



Analyse temps-fréquence et synthèse granulaire des bruits moteur diesel au ralenti: Application pour étude perceptive dans le contexte des scènes auditives.

Irina Zhekova

► To cite this version:

Irina Zhekova. Analyse temps-fréquence et synthèse granulaire des bruits moteur diesel au ralenti: Application pour étude perceptive dans le contexte des scènes auditives.. Acoustique [physics.class-ph]. Université de la Méditerranée - Aix-Marseille II, 2007. Français. NNT: . tel-00491452

HAL Id: tel-00491452

<https://theses.hal.science/tel-00491452>

Submitted on 11 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analyse temps-fréquence et synthèse granulaire des bruits moteur diesel au ralenti:

Application pour étude perceptive dans le contexte des scènes auditives

Thèse de doctorat présentée par

Irina Zhekova

en vue de l'obtention du grade

Docteur en Acoustique

Soutenue le 26 octobre 2007 au LMA devant un jury composé de

Philippe Guillemain
Richard Kronland-Martinet
Vincent Roussarie
Philippe Depalle
Laurent Simon
Etienne Parizet

Directeur de thèse
Co-directeur de thèse
Tuteur PSA
Rapporteur
Rapporteur
Examineur

Remerciements

« Celui qui se lance dans la ronde
doit danser jusqu'au bout »

Proverbe bulgare

Vous venez d'ouvrir la première page de mon manuscrit de thèse et je vous en remercie. Même si c'est amusant de cueillir les fruits de l'arbre de la connaissance pour soi-même, le partage donne un sens à tous les efforts. Je voudrais remercier ici les gens qui ont donné un sens à mon travail.

Ce travail n'aurait jamais pu voir le jour sans mon directeur de thèse, Philippe Guillemain, qui a cru en moi et ainsi a changé ma vie. Il m'a donné cette chance exceptionnelle de revenir en France, d'apprendre et d'évoluer. Je lui suis reconnaissante pour son vif et authentique esprit de recherche, sa bienveillance et son soutien jusqu'au bout, ainsi que pour les heures passées dans la cuisine des sons. En me laissant la liberté de mûrir et la responsabilité de bâtir mes propres idées, il m'a également appris une leçon de vie.

J'ai eu la chance d'avoir à mes côtés la double expertise de Vincent Roussarie. Il a été mon guide dans le monde de l'automobile et l'univers de la perception sonore. Toujours soucieux de me transmettre son savoir, il a supervisé avec attention et passion les tests perceptifs. Son intuition et ses encouragements m'ont été extrêmement bénéfiques. Je le remercie de m'avoir accompagnée pendant ce long chemin, pas toujours facile, et de m'avoir fait rencontrer les bonnes personnes au bon moment. Je remercie Philippe et Vincent pour les bons moments passés ensemble et pour leur générosité et bonne humeur.

Je suis très honorée par mes rapporteurs, Philippe Depalle et Laurent Simon. Je les remercie, avec beaucoup de confiance et respect pour leur qualités scientifiques, d'avoir accepté d'évaluer mon travail. Leur patience et leur sens du détail, ainsi que la justesse de leurs remarques constructives m'ont permis d'améliorer ce manuscrit. Je remercie également Etienne Parizet pour sa rigueur scientifique.

Je remercie Richard Kronland-Martinet et Anne Bardot, deux chefs d'équipe ouverts et compétents, de m'avoir accordé(e) (à) la symphonie humaine de leur équipes respectives, S2M (Signaux Sonores et Musicaux) et PEFH (Perception et Facteurs Humains). Je remercie aussi les chercheurs et les spécialistes rencontrés au Laboratoire de Mécanique et d'Acoustique à Marseille et à Peugeot Citroën à en région parisienne qui m'ont aidé à différentes étapes de mon travail.

J'ai eu le bonheur de faire partie de S2M, une équipe formidable avec qui j'ai vécu dans une perpétuelle découverte de la science et de la musique. Je remercie particulièrement Gregory Pallone pour les conversations enrichissantes et Mitsuko Aramaki qui m'a soutenu dans des moments difficiles. Un merci spécial à Thierry Voinier, mon sauveur des soucis informatiques à tout instant. Un grand merci à tous mes collègues, car chacun d'eux m'a aidé à sa manière. Enfin, j'ai fait l'agréable expérience que la douce odeur de café, la fumée du barbeque et le piano fêru à 4 mains illuminent à jamais l'esprit du chercheur.

Ma chance a été double, car mon équipe PEFH m'a manifesté une solidarité et bonne entente à toute épreuve. Mon "baptême de feu" sur les pistes d'essai de la Ferté Vidame a été rendu possible grâce à Mickaël Sauvage, un grand merci pour son efficacité et son amitié bretonne. Je remercie de tout cœur Florent Richard, pour sa complicité et son soutien constant, il a réfléchi sur toutes mes questions pour trouver une réponse et a cherché à résoudre tous mes problèmes avec beaucoup de compétence. C'est un plaisir de l'avoir comme collègue et ami. Je remercie aussi Olivier Sauvage qui s'est intéressé de près à ma thèse, pour sa disponibilité et son expertise sur la partie concernant l'acoustique automobile.

Une des périodes les plus délicate de ma thèse fut la réalisation des tests perceptifs. Un grand merci à Marie-Céline Bezat et Thibaut Jeandel, ainsi qu'à Yann et Alexis, et tous mes collègues de PEFH qui m'ont aidé dans cette tâche à la fois lourde et passionnante. Un merci à Olivier Lescop pour les logiciels de tests et les tests de la logistique. Je remercie aussi Pauline Faye et Carine Egoroff pour leurs conseils d'expert. Un grand merci à celle qui m'a réconciliée avec les statistiques, Emmanuelle Diaz, son aide lors du traitement des données m'a été indispensable et précieuse. Enfin, je veux remercier tous mes auditeurs qui ont passé du temps à écouter et évaluer les bruits au ralenti, et qui ont même pris du plaisir (!), sans eux la dernière partie de ma thèse n'aurait pas pu exister. Un clin d'œil à Brigitte.

Enfin, pour que mon aventure perceptive soit complète, j'ai eu le droit, en plus de la chaleur humaine, de petits fours avec du jus d'orange, des Tontons Flingueurs et de Raymond Davos, de pots d'arrivée et de départ, de pauses kiss-cool signées Steph, d'une bonne dose de rire tous les jours, et bien plus encore.

Ce manuscrit a été rédigé sans aucune correction automatique d'orthographe. Ce fut un défi pour moi et certainement pas très commode pour mon directeur de thèse et mes correcteurs, mais je me suis efforcée à donner le meilleur de moi-même et j'ai même eu raison dans quelques cas pointus. Cette thèse a été l'expérience ultime de mon français, une langue si riche, qu'elle représente une recherche en soi.

J'ai passé avec elle des mois à temps complet, je remercie donc l'administration française de m'avoir éclairée sur deux questions fondamentales : pourquoi faire compliqué alors qu'on peut faire simple ? (pour apprendre à relativiser dans la vie ..) et pourquoi la double épaisseur du papier toilette ? (car le deuxième exemplaire reste pour le dossier ..). Plus sérieusement, je remercie Laurence et Michèle d'avoir fait ma vie administrative plus agréable.

Un grand merci à la sagesse et l'efficacité hors pair de Dr Dubois qui a veillé, tel un ange gardien, sur mon ASICS (Anima Sana In Corpore Sano), et ce fut une chance pour moi de découvrir F.M. Alexander et Dr Bach, dont la science des fleurs fait des miracles.

Si je suis revenue en France après mon Erasmus, c'est aussi grâce et pour les gens merveilleux que j'ai pu rencontrer. Je remercie de tout mon cœur mes amis de cette auberge française pour leur amitié sans faille et pour les magnifiques coins de la France qu'ils m'ont fait découvrir. Ricoo, Alex, Ben, Manu B., Manu S., Manu Y. et Caro L., Julien R., Math, Caro LYK et Guigui, Benj et Caro, Kathrin et Mehdi, Aloïs, Debo, Nono, Patricia et toute la bande parisienne, Dimana, Gabriel de (not) Toulouse (but to win !), Sylvia et Christophe, et un merci très très spécial à Snezh, ainsi qu'à Geoffroy et sa famille.

Je remercie ma mère Kostadinka et mon frère Alexander, ainsi que ma grand-mère Raina et mon cousin Julian pour leur amour et soutien inconditionnel. Merci à mes fidèles amis, Snej, Steli et Emo de Varna, Kali de Sofia, Kati de Munich, Rosi de New York et à ma famille qui m'encourage et m'envoie depuis toujours des ondes positives de la Bulgarie. J'ai une pensée profonde à la mémoire de mon père Vladimir, à qui je dois cette double passion pour la science et la musique, et à celle de mon grand-père Ivan qui m'a transmis son amour pour la langue française. Je sais qu'ils auraient été fiers de moi.

A la fin, mais pas en dernier, je veux remercier Charlie pour toute la joie de vivre qu'il me donne au quotidien. Son allégresse, son amour et son humour sont un délice pour moi. Un grand merci aussi à sa formidable famille pour leur coup de pouce lors de la dernière ligne droite.

MERCI à la vie.

Irina Zhekova

Table des matières

Introduction	11
Contexte automobile	11
Contexte scientifique	11
Objectifs et structure de la thèse	12
Qu'est-ce que le bruit ?	13
I Le bruit moteur : origines et perception	15
1 Notions de base sur le moteur	19
1.1 Vue d'ensemble du moteur	19
1.2 Cycle moteur à 4 temps	20
1.3 Diagramme des événements	21
1.4 Spectre du bruit moteur	22
2 Acoustique moteur	23
2.1 Excitations primaires	23
2.1.1 Mouvements du moteur	23
2.1.2 Forces d'explosion	23
2.1.3 Forces d'inertie	24
2.1.4 Equilibrage des forces	25
2.1.5 Pressions pulsées	26
2.1.6 Résumé des sources primaires	26
2.2 Sources secondaires	26
2.2.1 Distribution	26
2.2.2 Injection	27
2.2.3 Admission	27
2.2.4 Echappement	28
2.2.5 Transmission	28
2.3 Autres sources	28
2.3.1 Alternateur	28
2.3.2 Pompe de direction assistée	29
2.3.3 Boîte de vitesse	29
2.4 Identification des sources sonores	29
3 Acoustique véhicule	33
3.1 Bruit de roulement	33
3.1.1 Processus mécanique	34
3.1.2 Résonance d'air	35
3.1.3 Propagation	35
3.1.4 Friction	36
3.1.5 Exemple de bruit de roulement	36
3.2 Bruit aérodynamique	36
3.2.1 Principaux mécanismes	36
3.2.2 Modélisation physique	37
3.2.3 Prédiction du bruit aérodynamique	38
3.3 Chemins de transfert	39
3.3.1 Transfert solidien	39
3.3.2 Transfert aérien	40
3.4 Ambiance sonore dans l'habitacle	41

4	Qualité de bruit automobile	43
4.1	Perception et qualité sonore	43
4.1.1	Démarche perceptive	43
4.1.2	Objectifs de la qualité sonore	44
4.1.3	Aspects de la qualité sonore automobile	44
4.2	Indicateurs psychoacoustiques	46
4.2.1	Sonie ("loudness")	46
4.2.2	Hauteur ("pitch")	48
4.2.3	Acuité ("sharpness")	49
4.2.4	Timbre	49
4.2.5	Rugosité ("roughness")	50
4.2.6	Fluctuation ("fluctuation strength")	52
4.3	Méthodologie des tests perceptifs	52
4.3.1	Test d'estimation de grandeur	53
4.3.2	Test d'évaluation	53
4.3.3	Test d'évaluation comparée	53
4.3.4	Test de différentiel sémantique	54
4.3.5	Test de comparaison par paires	54
4.3.6	Test de catégorisation (tri)	55
4.3.7	Test de description verbale	55
5	Moteur diesel au ralenti	57
5.1	La combustion	57
5.1.1	Pression dans le cylindre	57
5.1.2	Voies de propagation	57
5.1.3	Importance de l'injection	58
5.1.4	Qualité du bruit de combustion	59
5.2	Etude perceptive du ralenti	61
5.2.1	Sons impulsifs	61
5.2.2	Notion d'irrégularité	61
5.2.3	Exemple de configuration rythmique	62
5.2.4	Approche temps-fréquence du bruit au ralenti	64
II	Analyse et synthèse temps-fréquence	67
6	Les représentations temps-fréquence	71
6.1	Quelques définitions	71
6.2	Transformées de Gabor et de Fourier à court terme	73
6.2.1	Définition	73
6.2.2	Fréquence locale et retard local	74
6.2.3	Concept atomique	75
6.2.4	Principe d'incertitude de Heisenberg-Gabor	76
6.2.5	Echantillonnage du schéma de Gabor	76
6.3	Transformée en ondelettes	77
6.3.1	Définition et propriétés	78
6.3.2	Analyse multirésolution	80
6.3.3	Types d'ondelettes	81
6.4	Distribution de Wigner-Ville	81
6.4.1	Définition et propriétés	81
6.4.2	Interférences	83
6.4.3	Généralisation - classe de Cohen	84
6.5	Représentation en échelle perceptive	85
6.5.1	Echelle de Bark	85
6.5.2	ERB - Equivalent Rectangular Bandwidth	87
6.5.3	Filtres gammatones	87
6.5.4	Méthodes temps-fréquence intégrant des critères perceptifs	88
6.6	Applications pour les bruits moteur	88
6.6.1	Transformée de Gabor	89
6.6.2	Transformée en ondelettes	89

6.6.3	Echelle perceptive	90
6.6.4	Distribution de Wigner-Ville	91
7	La réallocation - une concentration énergétique	93
7.1	Opérateurs de la réallocation	93
7.1.1	Définition originale	93
7.1.2	Généralisations	95
7.1.3	Réallocation différentielle	96
7.2	Interprétation de la réallocation	97
7.2.1	Lien avec la fréquence instantanée	97
7.2.2	Lien avec les méthodes "arête et squelette"	98
7.3	Applications et exemples	99
8	Algorithmes de décomposition	101
8.1	Problématique	101
8.2	Approximations adaptatives	102
8.2.1	Optimalité	102
8.2.2	Critères de parcimonie	102
8.3	Matching poursuit	103
8.3.1	Algorithme	103
8.3.2	Atomes réels	104
8.4	Extensions du matching poursuit	104
8.4.1	Atomes chirps multi-échelle	104
8.4.2	Ridge poursuit	105
8.5	Exemples	106
9	Synthèse sonore	109
9.1	Classification des méthodes	109
9.1.1	Méthodes "de signal"	109
9.1.2	Méthodes physiques	110
9.1.3	Concept d'analyse-synthèse	111
9.2	Synthèse additive	112
9.2.1	Démarche	112
9.2.2	Extraction des paramètres	113
9.2.3	Synthèse additive des bruits moteur	114
9.3	Synthèse par modèles physiques	116
9.3.1	Instruments de musique et automobile	116
9.3.2	Modélisation physique des bruits de moteur	117
9.4	Synthèse granulaire	119
9.4.1	Classification des méthodes	119
9.4.2	Paramètres du grain	120
9.4.3	Phénomènes psychoacoustiques	121
10	L'analyse-synthèse du ralenti moteur	125
10.1	Bases de la méthode d'analyse-synthèse	125
10.1.1	Modèle atomique du signal	125
10.1.2	Schéma d'analyse-synthèse	126
10.2	Estimation des paramètres	128
10.2.1	Transformée de Gabor d'un atome non chirpé	128
10.2.2	Transformée de Gabor d'un atome chirp	129
10.2.3	Estimateurs pour β_k et σ_k	130
10.2.4	Opérateurs de réallocation temps-fréquence	131
10.2.5	Exemples de représentation	132
10.3	Algorithme d'estimation des paramètres	135
10.3.1	Maxima locaux	135
10.3.2	Critère de distance	136
10.3.3	Rapport entre σ et σ_k	137
10.3.4	Algorithme itératif	137
10.3.5	Estimation de l'amplitude A_k	137
10.3.6	Exemples d'estimation	138

10.4	Synthèse granulaire	143
10.5	Analyse-synthèse des bruits moteur	144
10.5.1	Extraction d'un cycle	144
10.5.2	Estimation des paramètres	146
10.5.3	Algorithme itératif	149
10.5.4	Synthèse	150
III	Perception des bruits moteur au ralenti	155
	Résumé de la troisième partie	157
11	Analyse des scènes auditives	159
11.1	Principes de l'analyse des scènes auditives	159
11.1.1	Traitement de l'information auditive	159
11.1.2	Théorie de Gestalt	160
11.1.3	Perception des séquences auditives	161
11.1.4	Attention auditive	162
11.2	Le bruit de moteur dans le contexte des scènes auditives	163
11.2.1	Bruit moteur et flux auditifs	163
11.2.2	Structure du bruit moteur	164
11.2.3	Notion de rythme	164
11.3	Exemples rythmiques au ralenti	165
11.3.1	Filtrage coupe-bande	165
11.3.2	Soustraction des atomes	165
11.3.3	Soustraction d'une zone temps-fréquence	166
11.3.4	Soustraction de plusieurs zones	167
12	Tests perceptifs pour les bruits au ralenti	169
12.1	Cycle moteur : étude en quarts	169
12.1.1	Découpage du cycle en quarts	169
12.1.2	Sonie spécifique des quarts	170
12.2	Concept de zones temps-fréquence	171
12.2.1	Définition	171
12.2.2	Description des zones	172
12.2.3	Construction des sons	173
12.2.4	Niveau relatif des zones	175
12.2.5	Effet des zones sur le niveau global	176
12.2.6	Test d'égalisation	177
12.3	Outils statistiques	178
12.3.1	Plan d'expérience	178
12.3.2	Facteurs et modalités	179
12.3.3	Optimalité d'un plan	179
12.3.4	Analyse multidimensionnelle (MDS)	180
12.3.5	Analyse en composantes principales (ACP)	181
12.4	Test de catégorisation	183
12.4.1	Protocole des tests	183
12.4.2	Présentation des dimensions	184
12.4.3	Effets des zones sur les dimensions	186
12.4.4	Interprétation des dimensions	187
12.4.5	Verbalisation	190
12.5	Test de préférence	191
12.5.1	Cartographie des préférences	191
12.5.2	Discussion des résultats	193
	Apport et perspectives	197
	Annexes	199
A	Les constituants du moteur	199

B Mesures vibroacoustiques	201
B.1 Introduction	201
B.2 Objectifs des mesures	201
B.3 Matériel	201
B.4 Positionnement des capteurs	202
B.4.1 Compartiment moteur	202
B.4.2 Roues	204
B.4.3 Echappement	204
B.4.4 Habitacle	204
B.5 Description des mesures	205
B.5.1 Mesures en statique	205
B.5.2 Mesures dynamiques	205
B.6 Traitement des données	206
C Liste des exemples sonores	207
Bibliographie	213

Introduction

Qu'est-ce que évoque naturellement dans notre esprit une voiture ? Notre imagination peut tourner autour de plusieurs de ses aspects : la puissance, la vitesse, la forme, la couleur, le style, le confort, etc. Cependant, la réaction innocente d'un enfant se résume souvent à un seul et incontestable "vroum-vroum". Notre première image imitative de la voiture est donc associée au bruit. Cette habitude inconsciente de se référer au bruit se manifestera plus tard dans toutes les situations de conduite comme le démarrage, le changement de rapport de vitesse, l'accélération et la décélération, le dépassement, l'avertissement sonore. Le bruit automobile est porteur d'information à la fois pour les conducteurs et pour les piétons. Le bruit peut être une cause de gêne auditive, mais également une source de plaisir. Imaginer une voiture sans bruit ? C'est bien là un des problèmes des voitures électriques.

Contexte automobile

La présente thèse se propose d'étudier les bruits moteur diesel au ralenti. Ce sujet très précis s'inscrit dans trois thématiques plus génériques : le bruit automobile, le bruit diesel et le bruit au ralenti.

Bruit automobile Le bruit représente un aspect indissociable de l'image de la voiture et contribue grandement à sa qualité globale. Le bruit dans l'habitacle doit donc satisfaire les exigences du client en termes de confort et d'identité sonore. "Une grosse berline ne doit pas sonner comme un tracteur" ¹, et une voiture sportive doit laisser entendre le caractère puissant du moteur. Cependant, même si le ressenti du client est très important, le contrôle prioritaire des concepteurs automobile est orienté vers le bruit émis à l'extérieur de la voiture. Ce contrôle est imposé par des lois gouvernementales et internationales afin de préserver l'environnement. Les lois limitent le niveau sonore émis à l'extérieur de la voiture dans l'objectif de restreindre la nuisance sonore du trafic routier.

Moteurs diesel Le moteur diesel fonctionne par autoinflammation des gaz et provoque une excitation vibro-acoustique beaucoup plus importante qu'un moteur à essence. Toutefois, dans les dernières années, les moteurs diesel se sont imposés sur le marché automobile en dépassant largement en chiffre de vente les voitures à essence [1]. La popularité des voitures diesel est due au développement des technologies comme l'injection directe "common rail" et l'injection pilote, ainsi que l'innovation en termes de réduction de la pollution avec le filtre à particules mis en place par le groupe PSA Peugeot Citroën. Un progrès considérable est également fait en vue de la réduction du bruit des moteurs diesel et le terme péjoratif "vieux diesel" appartient désormais au passé. Toutefois, avec l'augmentation de la puissance et du rendement, la structure des moteurs diesel subit des sollicitations de plus en plus importantes. La problématique de l'émission vibro-acoustique des diesels maintient donc toute son actualité.

Ralenti diesel Le ralenti est une situation particulière où le moteur tourne alors que la voiture demeure à l'arrêt. Le bruit au ralenti exprime donc la seule contribution du moteur, puisque les bruits dûs au roulement et d'écoulement d'air ne sont pas présents. Le bruit au ralenti pose un problème acoustique car il fait partie des bruits impulsifs répétitifs, souvent susceptibles de provoquer la gêne auditive. Malgré les améliorations apportées, le bruit au ralenti des diesels est nettement plus fort que celui des moteurs à essence. Le bruit au ralenti accompagne les situations souvent non désirées d'attente du conducteur, par exemple au feu rouge ou lors des embouteillages, et ces facteurs peuvent davantage augmenter la gêne.

Contexte scientifique

Le contexte scientifique de la thèse comprend trois domaines : les méthodes d'analyse temps-fréquence du signal, les méthodes de synthèse et l'analyse des scènes auditives. Concrètement sont choisis trois outils principaux : l'analyse temps-fréquence de Gabor, la synthèse granulaire et la théorie des flux auditifs.

Méthodes temps-fréquence Les bruits moteur au ralenti possèdent une structure à la fois temporelle et spectrale. L'outil d'analyse le plus adapté est donc à trouver parmi les méthodes temps-fréquence

¹Communication non formelle avec Vincent Roussarie, responsable du groupe perception auditive chez PSA

ou temps-échelle qui traitent le bruit moteur d'un point de vue du traitement de signal. D'une manière différente, les méthodes numériques comme les éléments finis (FEM) effectuent la modélisation physique du comportement vibro-acoustique de la structure du moteur. Notre méthode utilise les outils d'analyse du traitement du signal, mais adopte une approche phénoménologique envers le bruit moteur, consistant à garder un lien avec les événements temporels dans le cycle moteur. L'analyse d'un cycle moteur vise à extraire l'information essentielle du signal afin de le reconstruire à l'aide de signaux localisés en temps et en fréquence, appelés des atomes. Cette démarche emprunte également d'autres outils comme la réallocation temps-fréquence et les décompositions atomiques parcimonieuses. L'intérêt d'une représentation temps-fréquence du signal réside dans la facilité de manipulation des composantes du signal de synthèse.

Méthodes de synthèse D'une manière similaire aux méthodes d'analyse, la synthèse des signaux peut être faite à l'aide de la modélisation physique des sources acoustiques ou à l'aide de la modélisation des propriétés du signal sonore. Les méthodes physiques trouvent une application dans le contexte automobile, mais elle sont souvent très coûteuses en temps de calcul. Parmi les méthodes du signal, la plus utilisée dans la synthèse des bruits moteur est la synthèse additive. Elle modélise le contenu spectral du signal et offre un réalisme très satisfaisant des bruits à un régime moyen et élevé. En revanche, elle semble moins adaptée à la structure plutôt temporelle et impulsive des bruits au ralenti. Afin de combiner une simplicité de synthèse et de s'adapter à la structure des bruits au ralenti, notre choix est porté vers la synthèse granulaire. La synthèse granulaire est un outil complémentaire à l'analyse temps-fréquence puisqu'elle reconstruit le signal sous la forme d'une somme d'atomes temps-fréquence.

Perception des bruits L'étude perceptive des bruits automobile est motivée par le souci de satisfaire le client en termes de qualité sonore. Les propriétés subjectives des sons sont souvent étudiées à l'aide des indicateurs psychoacoustiques associés au niveau sonore, la hauteur ou plus globalement au timbre des sons. Toutefois, les particularités des bruits moteur exigent une adaptation de la définition de ces indicateurs. Un domaine récent de la perception auditive comprend la théorie des scènes auditives, traitant les règles d'organisation de séquences auditives. De multiples exemples des séquences de sons permettent de faire une équivalence avec les bruits au ralenti. Cette approche originale développée dans la thèse propose une nouvelle manière d'étudier les propriétés perceptives des bruits impulsifs répétitifs.

Objectifs et structure de la thèse

Dans cette thèse je me propose de rassembler des connaissances de trois domaines différents. Ce sont l'univers de l'acoustique automobile, les méthodes de d'analyse-synthèse des signaux et l'étude des phénomènes psychoacoustiques.

Motivation La motivation majeure de cette thèse est sa pluridisciplinarité. Les domaines des méthodes d'analyse-synthèse temps-fréquence et de la perception auditive sont réunies au service d'une problématique commune, l'acoustique automobile. Chacun de ces trois domaines représente à lui seul un vaste champ de recherche. La motivation n'est donc pas dans le sens d'aller en profondeur dans l'exploration de chaque domaine, mais dans la construction d'un fil conducteur les reliant dans une approche originale.

Objectif de la thèse La thèse a pour objectif principal de proposer une nouvelle méthode d'analyse temps-fréquence et synthèse granulaire, adaptée aux bruits moteur au ralenti. Cette méthode permet le découpage schématique du cycle moteur en zones temps-fréquence et l'interprétation du bruit moteur au ralenti dans le contexte des flux auditifs. Les tests perceptifs avec les bruits de synthèse se proposent de confirmer l'hypothèse de la formation des flux auditifs et de déterminer les principales dimensions perceptives des bruits au ralenti. Un intérêt très particulier est réservé à l'aspect rythmique des bruits.

Structure de la thèse La thèse est organisée en trois parties, consacrées respectivement à l'acoustique automobile, les méthodes temps-fréquence et la perception auditive. La première partie présente le fonctionnement du moteur et ses multiples sources acoustiques, le bruit de roulement et d'écoulement, les mécanismes de transfert aérien et solidien vers l'habitacle, ainsi que la notion de qualité sonore et les particularités du bruit diesel au ralenti. La deuxième partie décrit les méthodes temps-fréquence et inclut les différentes représentations (Gabor, ondelettes, Wigner-Ville), ainsi que la réallocation et la problématique des décompositions atomiques parcimonieuses du signal. Un chapitre est également consacré aux techniques comme la synthèse par modèles physiques et la synthèse additive pour les bruits automobile, ainsi que la synthèse granulaire dans le contexte de la perception auditive. La contribution majeure dans

la deuxième partie est faite par la description de la nouvelle méthode d'analyse-synthèse et son application aux bruits au ralenti. La troisième partie expose la théorie des scènes auditives et leur analogie avec les bruits au ralenti. Cette partie présente le nouveau concept de découpage du cycle moteur en zones temps-fréquence, ainsi que l'organisation de tests perceptifs des bruits de synthèse. La structure de la thèse est résumée sur la figure suivante (**Fig. 1**), où les points essentiels sont mis en valeur.

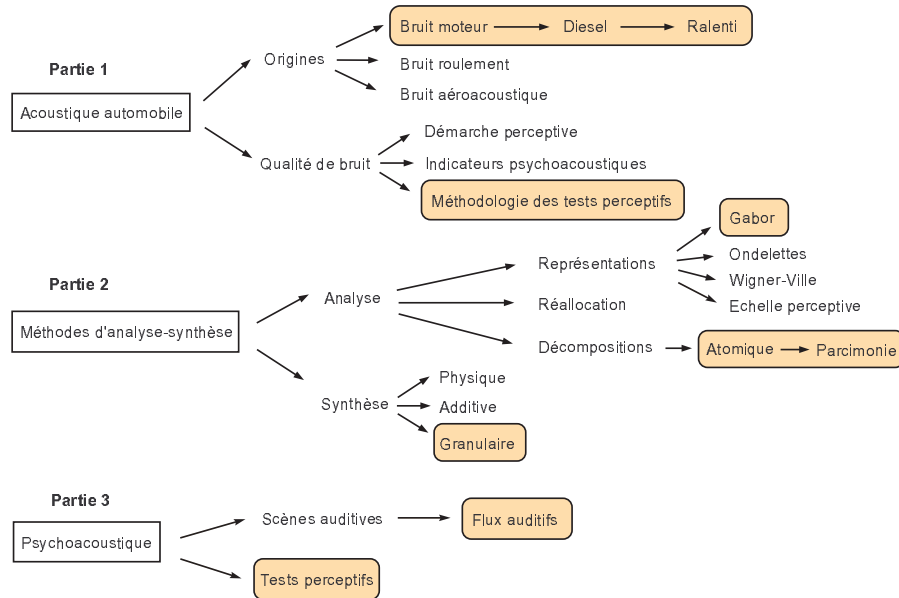


FIG. 1 – Représentation schématique de la thèse

Qu'est-ce que le bruit ?

La désignation "bruit" est quelque peu ambiguë et puisque tout au long de cet exposé on va parler de "bruit moteur", il est nécessaire de faire quelques considérations.

Définition du bruit La définition la plus commune du "bruit" le présente comme un "son indésirable", qui risque de provoquer une gêne auditive. Toutefois, l'aspect indésirable dépend fortement des conditions d'écoute où un même son peut être jugé agréable et désagréable. Une deuxième définition de "bruit" dans la taxonomie du traitement du signal le désigne comme son dont les fréquences apparaissent d'une manière aléatoire. On distingue le bruit blanc dont le spectre contient toutes les fréquences audibles, le bruit rose, le bruit marron etc. Cette définition exclut en effet l'aspect périodique du bruit. Le terme de "bruit moteur" semble ne pas être tout à fait en accord avec ces deux définitions. Le bruit moteur ne peut pas être jugé "indésirable", car il est une partie intégrante de l'image de l'automobile, et peut même procurer une sensation de bien-être et de puissance lors de la conduite. Le bruit moteur n'est pas purement aléatoire, mais possède une nature harmonique, générée par la rotation du vilebrequin.

Bruit et musique Les définitions semblent opposer le "bruit" à la "musique" sur des critères d'esthétisme sonore, en désignant avec "musique" la partie plaisante du monde auditif et avec "bruit" son aspect négatif. Une telle confrontation de nos jours n'a plus vraiment lieu, car la musique contemporaine explore de plus en plus des effets sonores du bruit en confirmant la pensée de Victor Hugo que la musique n'est qu'"un bruit qui réfléchit". Le compositeur italien Luigi Russolo crée même une taxonomie musicale des différents bruits communs dans son manifeste "L'Art des bruits" de 1913 [2]. Les compositeurs de la musique concrète sont également intéressés par la création des textures sonores expérimentales inouïes, qui bouleversent des habitudes mélodiques et harmoniques de la musique occidentale. Ces compositeurs se servent des sons désignés comme des bruits (bruit de réfrigérateur, de perceuse, de machine à écrire, etc.) en les intégrant dans des créations musicales. D'un tel point de vue, le bruit moteur peut être considéré même comme de la musique.

Convention Dans cette thèse "bruit moteur" va désigner "le son émis par le moteur", et le terme "bruit" sera dépourvu d'une connotation négative de gêne.

Première partie

Le bruit moteur : origines et perception

Résumé de la première partie

La première partie présente l'objet propre de la thèse - le bruit automobile. Pour se familiariser avec les particularités de cette classe de signaux très complexes, j'ai insisté sur trois points essentiels.

Le premier point concerne les principes de fonctionnement du moteur, et en particulier le cycle moteur à 4 temps. Par la suite, le moteur est décrit d'un point de vue acoustique, en détaillant les sources sonores et leur importance pour le bruit généré. Plus loin est présenté le bruit global du véhicule, avec les contributions du bruit de roulement et du bruit aérodynamique, ainsi que les chemins de transfert jusqu'à l'habitacle. Ces questions font l'objet des trois premiers chapitres : "Notion de base sur le moteur", "Acoustique moteur" et "Acoustique véhicule".

Le deuxième point essentiel traite la notion de qualité sonore des bruits moteur. Pour situer la problématique, il est nécessaire d'exposer les indicateurs psychoacoustiques comme la sonie, la hauteur, la rugosité etc, qui sont couramment utilisés pour qualifier les sons. Parfois, ces indicateurs génériques ne suffisent pas pour expliquer les particularités des bruits et notamment les bruits impulsifs répétitifs, c'est pourquoi quelques indicateurs spécifiques comme l'impulsivité ou la bruyance sont décrits. Enfin, une partie essentielle de l'étude sur la qualité sonore est occupée par les tests psychoacoustiques. Tous ces aspects de la qualité sonore sont inclus dans le chapitre "Qualité de bruit moteur".

Le troisième point important concerne le bruit moteur au ralenti et ses particularités et enjeux. Le chapitre, intitulé "Bruit moteur au ralenti", discute le processus de la combustion dans les moteurs diesel et son impact sur le bruit moteur. Le ralenti est présenté comme un régime moteur particulier, qui entraîne une perception plutôt temporelle du bruit. Dans ce contexte est introduite la classe des bruits impulsifs répétitifs, dont le ralenti moteur fait partie. Les bruits impulsifs représentent un grand intérêt d'un point de vue perceptif, car ils englobent la majorité des bruits industriels et sont étroitement liés à la gêne acoustique.

Notions de base sur le moteur

Le moteur à combustion interne transforme l'énergie chimique contenue dans le carburant, en énergie mécanique dégagée par l'explosion. L'énergie mécanique est utilisée par la suite pour le déplacement du véhicule. La dissipation d'une partie de cette énergie engendre des vibrations et un champ acoustique autour du véhicule. Le bruit émis par le véhicule dépend inévitablement du fonctionnement du moteur.

1.1 Vue d'ensemble du moteur

Le moteur est constitué de plusieurs cylindres confinant les explosions. Dans chaque cylindre coulisce un piston effectuant un mouvement alternatif. Ce mouvement de translation est transformé par l'intermédiaire de la bielle en mouvement de rotation du vilebrequin (**Fig. 1.1**). Reposant sur des paliers, le vilebrequin est lié à son tour au volant moteur et lui transmet le couple de forces, grâce auquel s'effectue la propulsion de l'automobile.

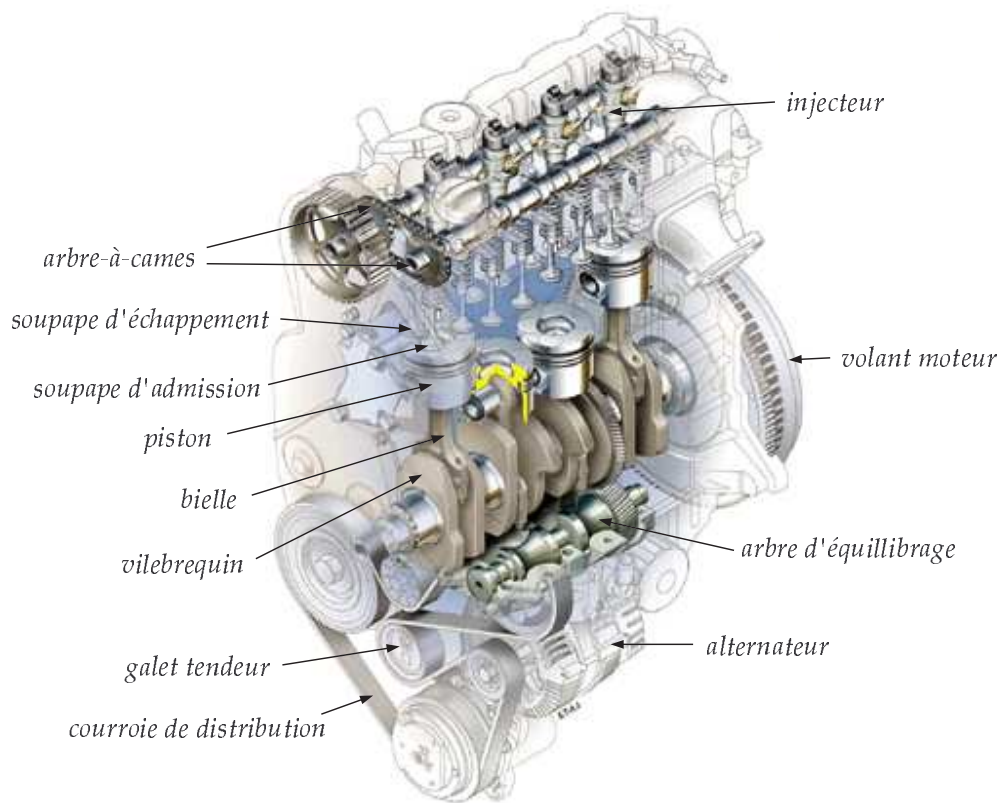


FIG. 1.1 – Vue d'ensemble d'un moteur (source image : [3])

Les explosions dans les cylindres sont déclenchées par l'étincelle des bougies dans le cas d'un moteur à essence ou par l'auto-inflammation des gaz pour un diesel. Le carburant qui alimente la combustion est introduit dans le cylindre par l'injecteur. Les soupapes d'admission et d'échappement gèrent l'arrivée d'air et la sortie des gaz brûlés. Les soupapes sont gouvernées, à leur tour, par des arbres à cames tournant à une vitesse deux fois plus rapide que le vilebrequin. Le mécanisme rotatif est maintenu par une courroie ou une chaîne de distribution. Le fonctionnement du moteur à combustion suit le cycle moteur à 4 temps, décrit en détail ci-après.

1.2 Cycle moteur à 4 temps

Le cycle moteur à 4 temps a été inventé par Beau de Rochas en 1862. Le principe du cycle de Beau de Rochas fût repris par Etienne Lenoir en 1863 pour la réalisation du moteur à 4 temps. Indépendamment, ce principe fût déduit et implémenté en 1876 par Nikolaus Otto, considéré comme le co-inventeur du moteur à combustion interne. Otto constata, entre autres, que le rendement réel du cycle moteur fût inférieur au celui du cycle théorique. Plus tard, Rudolf Diesel développa en 1893 l'idée d'un moteur à allumage par compression, le moteur diesel.

Principe Le cycle moteur à 4 temps est marqué par plusieurs événements physiques se succédant dans le temps. Ces événements sont regroupés en quatre phases (temps) : admission, compression, combustion (et détente) et échappement. Le rôle principal occupe le vilebrequin, car il gouverne les cycles de tous les éléments entraînés lors de sa rotation. Le cycle moteur se produit sur 2 tours de vilebrequin. Par ailleurs, chaque cylindre du moteur obéit au même schéma de fonctionnement. Ce schéma se décompose en quatre temps (Fig. 1.2).

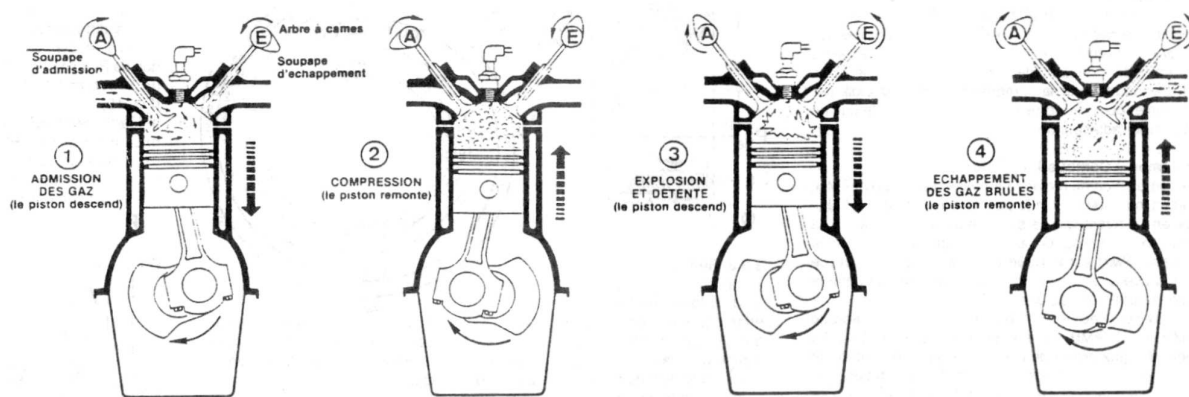


FIG. 1.2 – Les phases du cycle moteur à quatre temps (source [4])

1^{er} temps = Admission Le cycle moteur commence par l'admission. Le début de ce temps est marqué par le mouvement du piston, du point mort haut (PMH) au point mort bas (PMB), précédé juste avant de l'ouverture de la soupape d'admission. La course descendante du piston provoque une dépression dans le cylindre, qui facilite l'aspiration du mélange air/essence (pour les moteurs à essence) ou air (pour les diesels). Une fois le PMB atteint, la soupape d'admission se referme aussitôt. A ce moment, la pression des gaz dans le cylindre est égale à la pression atmosphérique.

2^e temps = Compression Le piston remonte du PMB au PMH et provoque ainsi la compression des gaz dans le cylindre. Le taux de compression devient très important pour les moteurs diesel, et il va permettre l'auto-inflammation de mélange air/gazole. Le vilebrequin termine son premier tour.

3^e temps = Combustion, Détente Pendant cette phase, au moment où le piston se trouve au PMH, les gaz explosent. Pour un moteur à essence, le mélange est enflammé par l'étincelle d'allumage des bougies, fixées sur la partie supérieure des cylindres. Pour un moteur diesel, l'explosion est due à la pression beaucoup trop importante des gaz qui génère l'auto-inflammation du combustible, injecté sous forte pression. Sous la poussée des gaz brûlés, le piston descend, en transmettant via la bielle, une force au vilebrequin. Juste avant le retour du piston au PMB, la soupape d'échappement s'ouvre.

4^e temps = Echappement La soupape d'échappement est maintenant ouverte afin de permettre la sortie des gaz. L'échappement est accompagné par la remontée du piston. La soupape d'échappement se ferme après l'arrivée du piston au PMH et crée ainsi une légère dépression dans le cylindre afin de faciliter l'entrée des gaz frais. Le vilebrequin termine son deuxième tour. Il conclut le cycle moteur à quatre temps. Il est essentiel de noter que le vilebrequin fait 2 tours par cycle moteur.

1.3 Diagramme des événements

Le schéma du cycle moteur à 4 temps décrit le fonctionnement d'un seul cylindre. Dans la présence de plusieurs cylindres, il y a une superposition, en décalé, de leurs propres cycles. En fait, tous les cylindres obéissent au même schéma, mais à un instant donné, ils se trouvent chacun dans une phase différente du cycle moteur. En règle générale, sur un cycle moteur, tous les cylindres passent à leur tour par une explosion. Le nombre d'explosions dans un cycle dépend du nombre de cylindres constituant le moteur.

Ordre d'allumage L'ordre d'allumage détermine le passage consécutif par une explosion de chaque cylindre. La configuration d'allumage est établie afin d'optimiser les forces et les couples d'un moteur. Elle va dépendre de la disposition des cylindres (le plus souvent en ligne ou en V), du nombre de cylindres (le plus souvent 4 ou 6) et de la cylindrée (en moyenne 2 l). La configuration la plus répandue pour un moteur à 4 cylindres en ligne est "1-3-4-2", le premier cylindre étant le plus près de la distribution.

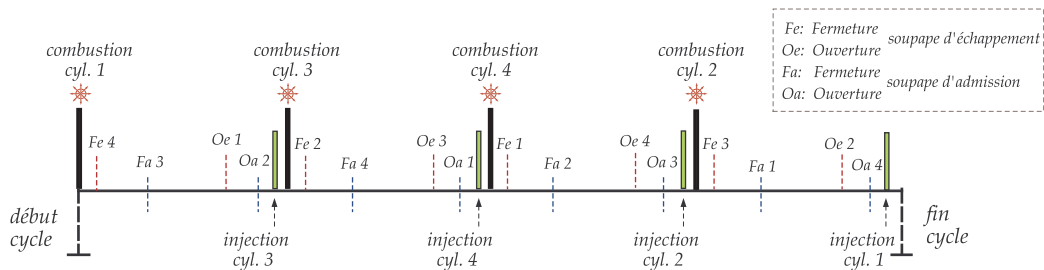


FIG. 1.3 – Le diagramme des événements lors du cycle d'un moteur à 4 cylindres

Superposition des cycles Le cycle moteur à 4 temps est exécuté avec un décalage dans chaque cylindre. Ce décalage superpose les cycles propres à chaque cylindre et détermine le diagramme global des événements qui se produisent lors d'un cycle moteur (**Fig. 1.3**). En fonction de l'ordre d'allumage des cylindres, on peut suivre à quel instant interviennent l'ouverture et la fermeture des soupapes d'admission et d'échappement, ainsi que l'injection et la combustion. Ces instants sont estimés en angle vilebrequin, où un cycle moteur correspond à 720° , i.e à 2 tours de vilebrequin. L'angle vilebrequin est une variable propre au fonctionnement du moteur. En fonction du régime moteur, elle peut être convertie en temps.

Différences entre les cylindres Malgré le fait que les cylindres fonctionnent selon un schéma identique, leurs explosions diffèrent légèrement. En effet, l'architecture spécifique à chaque moteur et les paramètres de réglage de la combustion (taux de compression, quantité de combustible injecté, etc.), provoquent des phénomènes aléatoires lors de l'explosion des gaz. Ces phénomènes rendent la combustion et le bruit émis différents de cylindre à cylindre, en attribuant une signature vibro-acoustique spécifique à chaque cylindre. Par ailleurs, la richesse sonore du bruit moteur augmente avec l'augmentation du nombre des cylindres. Ainsi, le monocylindre (moteur avec un seul cylindre) possède un bruit qui paraît trop "régulier" et "dur" à l'oreille, en comparaison avec un moteur 4 cylindres.

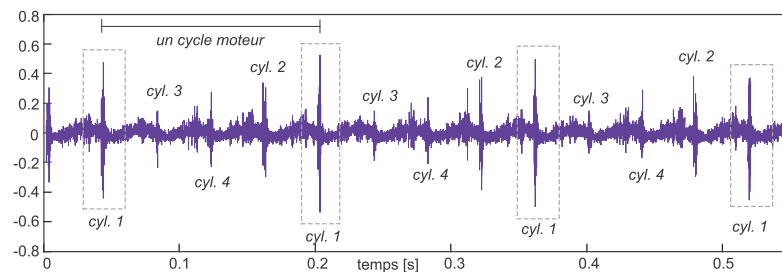


FIG. 1.4 – Signal accélérométrique à proximité du cylindre 1 d'un moteur à essence au ralenti

Différences cycle-à-cycle De même que les cylindres n'ont pas la même loi de combustion, les explosions d'un même cylindre varient d'un cycle à l'autre (**Fig. B.7** et **Piste 1** du CD joint). La variation entre les explosions d'un cycle à l'autre est associée à la nature non déterministe du processus de la combustion. Cette "irrégularité" cycle-à-cycle est davantage exprimée au ralenti qu'en haut régime.

1.4 Spectre du bruit moteur

Le fonctionnement du moteur est synchronisé avec la rotation du vilebrequin, exprimée en tours par minute (tr/min). Le spectre du bruit moteur va dépendre de la période des événements dans le cycle moteur et plus particulièrement des explosions. En effet, la richesse du spectre est le résultat d'un équilibre délicat entre les forces d'inertie et d'explosion. Ce processus, détaillé dans le chapitre suivant, est présenté ici dans une version simplifiée.

Périodes moteur Le fonctionnement du moteur crée notamment deux périodes liées au cycle moteur : la période globale du moteur et la période des explosions des cylindres. La période globale englobe le temps où tous les cylindres passent par une explosion. Elle ne dépend pas du nombre des cylindres, mais juste du régime moteur. En revanche la période propre aux explosions dépendra du nombre des cylindres. En notant N le régime moteur (en tr/min) et N_c le nombre de cylindres, on associe à ces périodes deux fréquences respectives :

$$\text{fréquence globale} = \frac{1}{2} \cdot \frac{N}{60} \quad [\text{Hz}] \quad \text{fréquence d'explosion} = \frac{N_c}{2} \cdot \frac{N}{60} \quad [\text{Hz}] \quad (1.1)$$

La fréquence globale pour un régime moteur $N = 3000$ tr/min est égale à 25 Hz. La fréquence d'explosion au même régime moteur va dépendre du nombre de cylindres et sera 100 Hz pour un moteur 4 cylindres et 150 Hz pour un moteur 6 cylindres. La durée d'un cycle moteur est inverse à la fréquence globale : pour un moteur tournant à 3000 tr/m, la durée du cycle est de 0,04 s, alors que la durée du cycle pour un régime au ralenti à 750 tr/m est de 0,160 s.

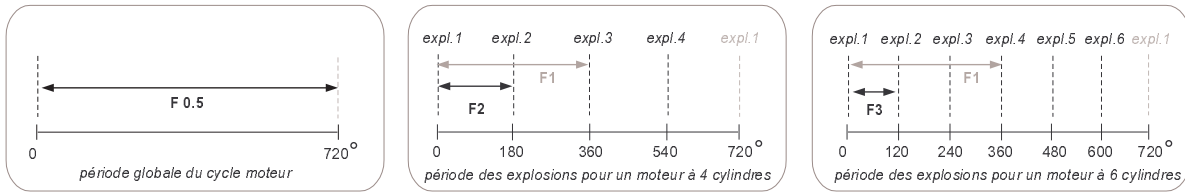


FIG. 1.5 – Période globale et période des explosions du cycle moteur

Fondamentale moteur En étudiant le spectre du bruit moteur, on utilise souvent le terme "fondamentale moteur". Pour des moteurs qui possèdent un allumage équidistant des cylindres, généralement la fondamentale moteur est liée à la fréquence des explosions. Ainsi, pour un moteur à 4 cylindres, la fondamentale moteur est F2, alors que pour un moteur à 6 cylindres, la fondamentale est F3 (Fig. 1.5). L'harmonique F1 correspond à un tour complet du vilebrequin. Enfin, on désigne la demi-harmonique F0.5, associée à la fréquence d'explosion d'un seul cylindre, dont la période occupe un cycle moteur, i.e 2 tours de vilebrequin. Pour un moteur à 4 cylindres, les plus marquées sont les harmoniques paires F4, F6, F8, etc, qui suivent la fondamentale moteur F2 (Fig. 1.6 et Piste 2) .

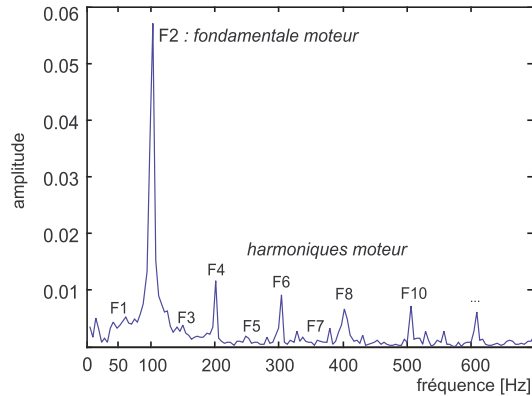


Fig. 1.6 – Spectre de bruit moteur à 3000 tr/min

Bandes de fréquences Le spectre du bruit moteur est souvent découpé d'une manière non formelle en trois bandes fréquentielles : basses [au-dessous de 500 Hz], moyennes [500-1500 Hz] et hautes fréquences [au-dessus de 1500 Hz]. Les limites fréquentielles sont données à titre indicatif, car elles varient selon les considérations. Les basses fréquences contiennent les premières harmoniques des forces et des couples d'explosion et inertie, ainsi que des pressions pulsées. Les moyennes fréquences comprennent des harmoniques de rang élevé des forces d'explosion et de pression, dont celles du carter cylindres. La principale contribution du bruit dans les hautes fréquences vient des chocs des parties mobiles du moteur sur la structure de la culasse et du carter cylindres. Un exemple de tels bruits est le claquement des soupapes, le bruit du jeu des pistons, ainsi que l'excitation la plus importante, la combustion.

Acoustique moteur

L'acoustique moteur concerne l'émission sonore du groupe moto-propulseur (GMP). Le GMP comprend le moteur même, ainsi que la boîte de vitesse. Les lignes d'échappement et d'admission sont considérées comme des éléments annexes du GMP. Le bruit moteur est le résultat du rayonnement acoustique de multiples sources différentes. Ces sources sont fortement corrélées et même si elles possèdent une signature sonore spécifique, leurs émissions se recouvrent spectralement, en rendant le bruit moteur très complexe. Ce chapitre introduit les principaux mécanismes et les types de sollicitations qui jouent un rôle important dans l'émission acoustique du moteur.

2.1 Excitations primaires

Le bruit moteur est le produit de trois types d'excitations primaires : des forces d'explosion, des forces d'inertie et des pressions pulsées. Les forces d'explosion et d'inertie sont liées à l'attelage mobile et à la distribution. Les forces d'explosion se créent dans le cylindre sous la pression variable des gaz et s'exercent sur le piston. Les forces d'inertie sont générées par l'attelage mobile et plus particulièrement les pièces rotatives du moteur comme la bielle, le piston, etc. Les pressions pulsées accompagnent l'entrée et la sortie des gaz à travers des soupapes d'admission et d'échappement et provoquent des résonances dans le conduit d'échappement. La description de ces forces, donnée par la suite, est essentiellement formulée dans le cas du monocylindre. Néanmoins, la présence de plusieurs cylindres et leur nombre, peut altérer l'équilibrage des forces. Par conséquent, les vibrations et le bruit émis se trouvent également modifiés.

2.1.1 Mouvements du moteur

La pression des gaz due à l'explosion génère une force longitudinale qui est transmise au piston et se transforme en couple au niveau du vilebrequin. Le moteur effectue des mouvements de translation et de rotation, sur les trois axes autour de son centre de gravité. On distingue au total six types de mouvements : roulis, galop, pilon, lacet, tamis transversal et tamis longitudinal (Fig. 2.1).

Le roulis est un mouvement de rotation accompli le long du vilebrequin, i.e autour de l'axe \vec{O}_x , et provoqué par les moments d'inertie et de pression. Le galop est une rotation autour de l'axe \vec{O}_y sollicitée par les forces d'inertie rotative et alternative. Le pilon est un mouvement de translation dans la direction de l'axe \vec{O}_z sollicité par les forces d'inertie rotative et alternative. Le lacet est un mouvement de rotation autour de l'axe des cylindres, \vec{O}_z , entraîné par la force d'inertie rotative. Le tamis transversal représente une translation dans la direction de l'axe \vec{O}_y , liée à la force d'inertie rotative. Le tamis longitudinal est une translation dans la direction de l'axe \vec{O}_x , liée à la force d'inertie rotative. Ces vibrations diverses sont propagées via les points d'attache du moteur, appelés encore des cales moteur, à travers la structure de la voiture.

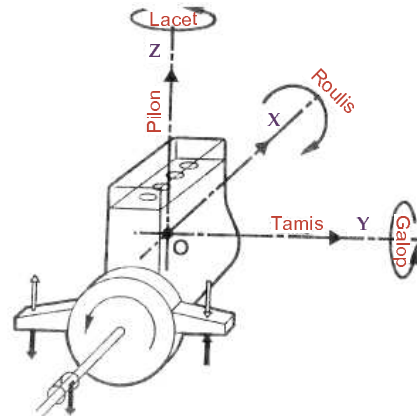


Fig. 2.1 – Les mouvements du moteur [4]

2.1.2 Forces d'explosion

La combustion transforme l'énergie contenue dans le carburant en deux types d'énergie. La première se présente sous la forme d'une chaleur perdue, évacuée par le circuit de refroidissement et par l'échappement. La deuxième forme est une énergie mécanique qui apparaît suite à la force de pression, engendrée par l'explosion. Cette force, exercée sur le piston, est transmise au vilebrequin via la bielle. Le processus crée ainsi un couple de forces au bout du vilebrequin qui est utilisé pour la propulsion du véhicule. On va décrire ici la genèse de ces forces dans le cas d'un monocylindre.

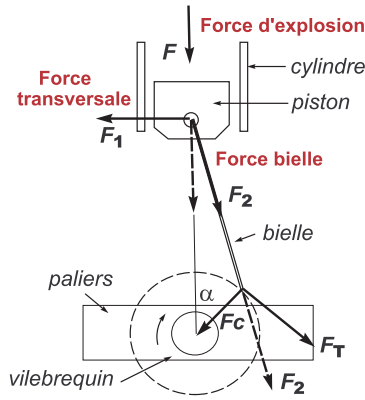


Fig. 2.2 – Forces d'explosions

Décomposition en série La force d'explosion s'exerçant au piston est périodique et décomposable en série de Fourier. Son moment M_{gaz} est aussi décomposable en série et s'écrit

$$M_{gaz} = M_0 + \sum_k A_k \cos \frac{\alpha}{2} + B_k \sin \frac{\alpha}{2} \quad (2.1)$$

où α est l'angle vilebrequin et M_0 est une constante. Le terme vibratoire est en $\frac{\alpha}{2}$ et il correspond à l'harmonique F0,5. En effet, pour un monocylindre, il y a une explosion tous les deux tours de vilebrequin. Dans la formule (2.1), les termes A_k et B_k ne dépendent pas du régime moteur, mais de la pression maximale mesurée lors d'un cycle. Deux paramètres influent le moment des forces d'explosion : l'état de charge et le type du moteur. Les forces d'explosion sont plus élevées en pleine charge qu'au ralenti, et également plus élevées pour un diesel que pour un moteur à essence.

2.1.3 Forces d'inertie

Les forces d'explosion sollicitent les pièces mobiles du moteur. Dans ce processus, la masse de la bielle force le piston à descendre et, en s'opposant à ce mouvement, il crée des forces d'inertie (Fig. 2.3). Les forces d'inertie se superposent aux forces d'explosion des gaz et font varier le couple moteur.

Type de forces et notation On distingue deux types de forces d'inertie : rotatives et alternatives. Ces forces sont décrites plus loin à l'aide de formules mathématiques, où sont utilisées les notations suivantes : M_{alt} est la masse du piston, γ_p est l'accélération du piston, λ est le rapport entre la longueur de la bielle L et le rayon manivelle R . Enfin, $\alpha = \frac{2\pi Nt}{60} = \omega t$ est l'angle de rotation du vilebrequin. Dans la notation relative au fonctionnement moteur N désigne le régime moteur en tours par minutes (tr/min). Les forces d'inertie engendrent des vibrations basses fréquences, en s'accroissant avec la vitesse de rotation du vilebrequin.

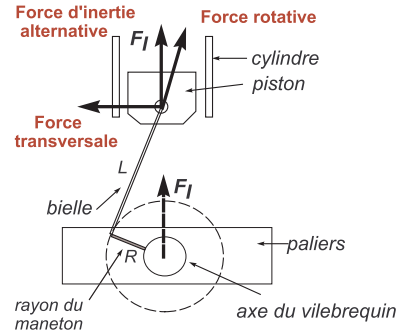


Fig. 2.3 – Forces d'inertie

Forces d'inertie alternatives Le mouvement alternatif du piston génère une force verticale, proportionnelle à l'accélération du piston, selon la loi classique $F_I = -M_{alt} \cdot \gamma_p$. Le développement en série de Fourier de l'accélération du piston est donné par la formule [5]

$$\gamma_p = -R\omega^2 \left[\cos \alpha + \frac{\cos 2\alpha}{\lambda} - \frac{\cos 4\alpha}{4\lambda^3} + \frac{\cos 6\alpha}{6\lambda^5} - \dots \right] \quad (2.2)$$

Le premier terme dans cette décomposition, $\frac{-R\omega^2 \cos 2\alpha}{\lambda}$, représente la fondamentale F1, désignée encore comme l'harmonique 1. Dans la décomposition (Eq. 2.2) restent seulement les harmoniques d'ordre pair (2F, 4F, 6F, ...). D'un point de vue vibro-acoustique, seules les deux premières harmoniques, F1 et F2, sont considérées importantes. Les harmoniques de rang élevé deviennent de plus en plus faibles, à cause du rapport fixe λ .

Forces d'inertie rotatives Les forces transversales génèrent un moment de la forme

$$M_{rot} = M_{alt} R^2 \omega^2 \left[\frac{-1}{4\lambda} \sin \alpha + \frac{1}{2} \sin 2\alpha + \frac{3}{4\lambda} \sin 3\alpha + \frac{1}{4\lambda^2} \sin 4\alpha \right]. \quad (2.3)$$

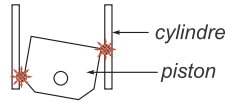


Fig. 2.4 – Basculement du piston

Ainsi, pour un 4 cylindres, la fondamentale est située à F2 et toutes les autres harmoniques sont ses multiples. La fondamentale s'étend de (25-200 Hz) en fonction du régime moteur de (750-6000 tr/min) et génère un bruit dit de bourdonnement.

2.1.4 Equilibrage des forces

L'équilibrage des forces d'explosion et d'inertie est primordial, non seulement pour le rendement du moteur, mais également pour ses caractéristiques acoustiques. Ici, je présente deux phénomènes importants et souvent cités par les motoristes et les acousticiens : le bourdonnement et l'acyclisme moteur et deux phénomènes moins couramment utilisés : le pilonnement et la rudesse moteur.

Bourdonnement Le bruit de bourdonnement ("booming") est le résultat de l'équilibrage des forces d'inertie et d'explosion. Pour un moteur 4 cylindres, les forces d'explosion et d'inertie sont en opposition de phase. Le bourdonnement est un bruit sourd qui s'apparente à un bruit de vol de bourdon. Il est propre au fonctionnement du moteur et est déterminé par la fondamentale moteur F2, comprise entre 20 et 200 Hz (**Fig.2.5** et **Piste 3**). Le bruit de bourdonnement dépend des paramètres comme l'architecture du moteur, les masses d'équilibrage, les suspensions et les fixations. Ce bruit peut être gênant et il est étudié dès la phase de conception du moteur.

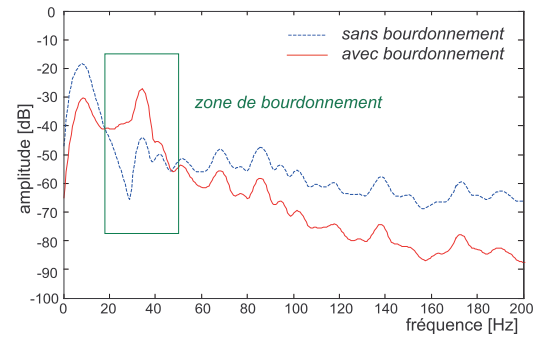


Fig. 2.5 – Spectre de bruit moteur à 1000 tr/min

Acyclisme moteur L'acyclisme moteur est la grandeur qui exprime la variation de la vitesse instantanée du vilebrequin. Le vilebrequin ne tourne pas à une vitesse constante, mais oscille à une vitesse moyenne, suivant la décomposition harmonique du moment du roulis. Ce moment est une combinaison du moment de roulis dû aux explosions et du moment dû aux inerties. Cette vibration excite les éléments tournants, le vilebrequin et l'attelage mobile et crée des jeux et des résonances. L'acyclisme moteur est un phénomène qui influence l'acoustique moteur en rendant le bruit moteur moins "régulier".

Pilonnement moteur Le pilonnement moteur a pour origines les vibrations du vilebrequin, couplé à la flexion du GMP, qui devient important en pleine charge et à un régime élevé. Le pilonnement est un bruit saccadé, répétitif et sourd qui occupe le spectre autour de 200 et 300 Hz et qui est dû, pour un moteur 4 cylindres, aux harmoniques F2.5, F3 et F3.5 (**Fig. 2.6** et **Piste 4**)

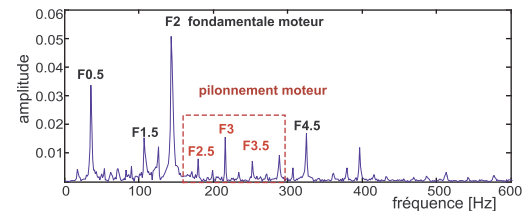


Fig. 2.6 – Pilonnement moteur à 4200 tr/min

Rudesse moteur La rudesse moteur est un phénomène dû aux résonances de flexion et de torsion du vilebrequin, excité par les harmoniques des forces d'explosion. La situation la plus propice où ce type de bruit se manifeste est la mise en charge et la pleine charge. La composante principale de la rudesse d'un moteur se situe entre 150 et 500 Hz, alors que certaines de ses harmoniques peuvent s'étendre de 300 à 1000 Hz (**Fig.2.7** et **Piste 5**).

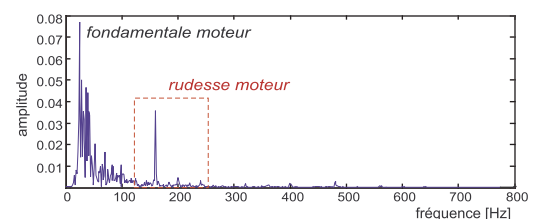


Fig. 2.7 – Rudesse moteur

2.1.5 Pressions pulsées

Les pressions pulsées représentent le troisième type d'excitation primaire du moteur. Elles sont liées au processus d'entrée et de sortie des gaz. Les pressions pulsées sont gérées au niveau des soupapes et font vibrer les parois des lignes d'admission et d'échappement, en provoquant aux extrémités de ces lignes des bruits appelés bruits de bouche. Etant synchronisées avec les explosions, les pressions pulsées contribuent dans le bruit moteur surtout aux composantes harmoniques paires, notamment F4, F6 et F8. [6]. En fonction de leur géométrie, les conduits d'admission et d'échappement représentent des sources secondaires non négligeables qui s'ajoutent au bruit moteur global.

2.1.6 Résumé des sources primaires

L'acoustique du moteur dépend fortement des excitations primaires dues aux forces d'inertie, d'explosion et les pressions pulsées. Ces excitations, liées à l'architecture et les réglages du moteur, rendent les émissions acoustiques différentes d'un moteur à l'autre. Parmi les phénomènes acoustiques importants on peut noter le basculement du piston et la combustion, deux contributions large bande difficilement séparables, car elles se produisent en un laps de temps très rapproché et couvrent une plage fréquentielle similaire [7]. Certaines études sur les moteurs diesel indiquent que le basculement du piston est prédominant à faible charge et à haut régime, alors que la combustion influence surtout en forte charge. Pour le moteur à essence les sollicitations vibratoires ne sont pas aussi fortes que pour un diesel.

2.2 Sources secondaires

Les sources secondaires apparaissent suite à une excitation liée aux sources primaires ou comme une excitation plutôt indépendante qui s'ajoute aux vibrations du moteur (**Fig. 2.8**).

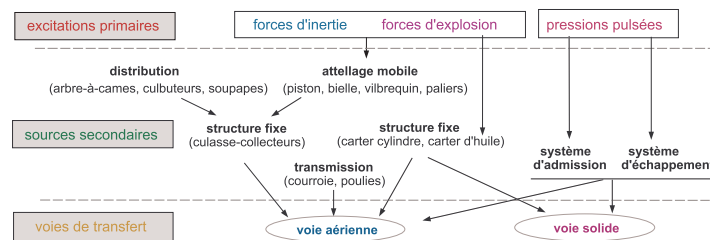


FIG. 2.8 – Schéma acoustique simplifié du moteur

J'ai regroupé ici les sources secondaires selon leur ressemblances fonctionnelles dans les groupes suivants : la distribution, l'injection, l'admission, l'échappement et la transmission.

2.2.1 Distribution

La distribution assure le mouvement des soupapes (**Fig. 2.9**). Elle se fait à l'aide d'un ou plusieurs arbres à cames, qui contrôlent l'ouverture et la fermeture des soupapes et tournent à une vitesse deux fois plus rapide que le vilebrequin. L'arbre à cames comporte des cames, chacune associée à une soupape. Lors de la rotation de l'arbre, les cames appuient sur les soupapes pour provoquer leur ouverture. En relâchant l'appui, les soupapes se ferment, repoussées par les ressorts. Ce mécanisme met en jeu deux types de sollicitations qui excitent la culasse du moteur.

Le premier type génère des chocs sur les sièges des soupapes lors de leur ouverture et provoque un bruit métallique lors de leur fermeture. Ces chocs ont une période globale de $F0.5$ par cylindre, car sur deux tours de vilebrequin, on compte une ouverture et une fermeture pour la soupape d'admission et de même pour la soupape d'échappement. Le deuxième type de sollicitation est associé aux arbres à cames qui subissent sur leurs paliers des efforts dynamiques créés par les forces d'inertie. Les vibrations mécaniques engendrées, sont accompagnées par un bruit, rayonné principalement par la culasse et le couvre culasse du moteur. Ce bruit dépend fortement du régime moteur et pas tellement de l'état de charge.

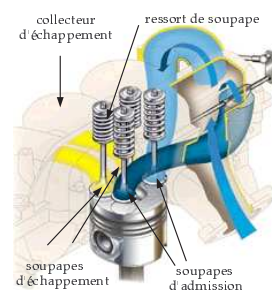


Fig. 2.9 – Soupapes [3]

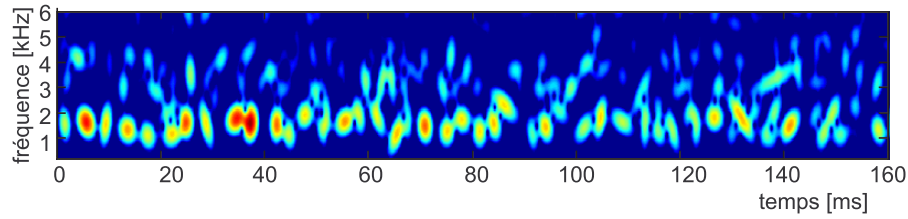


FIG. 2.10 – Spectrogramme d'un cycle moteur au ralenti, enregistré près des soupapes

Le bruit de distribution est très perceptible au ralenti et couvre le spectre de 1-4 kHz (**Fig. 2.10** et **Piste 6**). Dans cette bande fréquentielle se trouve aussi la contribution de l'injection, décrite ci-après. Sur la **Fig. 2.10** est représenté le spectrogramme temps-fréquence d'un bruit au ralenti enregistré près des soupapes. Par ailleurs, les méthodes temps-fréquence font l'objet de la deuxième partie de la thèse.

2.2.2 Injection

L'injection est un des paramètres clés de la combustion. On distingue deux types d'injection : directe et indirecte. Dans le premier cas, le combustible est injecté directement dans le cylindre ; dans le deuxième, l'injection se fait dans une pré-chambre, une partie auxiliaire de la chambre principale de combustion. L'injection directe représente plusieurs avantages pour les moteurs diesel, car elle permet une meilleure combustion et une réduction de la consommation.

Il existe principalement 3 techniques d'injection directe pour les moteurs diesel : l'injection classique, l'injection type rampe commune ("common-rail") et l'injecteur-pompe. Le système d'injection représente une pompe à haute pression qui introduit le carburant à l'aide des injecteurs. La pression d'injection est très élevée et est de l'ordre de 100 à 300 bars pour une injection directe classique, entre 1300 et 2000 bars pour une injection "common-rail" et plus de 2000 bars pour l'injecteur-pompe [8]. La forte pression pour l'injecteur-pompe permet de morceler très finement les particules du carburant (**Fig. 2.11**). Cette mécanique assure aussi l'injection pilote, qui consiste à introduire une faible quantité de carburant juste avant la combustion principale. Ce procédé permet une combustion plus douce et un bruit nettement plus faible.

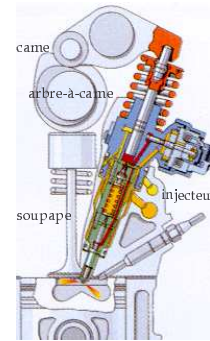


Fig. 2.11 – Injecteur-pompe [9]

Un bruit accompagne le début et la fin de l'injection, à 2-4 millisecondes d'écart en temps. La forte pression d'injection engendre un choc mécanique sur la structure, alors que la retombée de l'aiguille de l'injecteur provoque un bruit supplémentaire. Quand le capot du moteur est fermé, le bruit d'injection n'est pas très perceptible. Cependant, l'injection peut parfois représenter une source non négligeable de bruit. La qualité des injecteurs fait l'objet d'études acoustiques [10].

2.2.3 Admission

Le bruit d'admission vient de l'air, aspiré à l'entrée et pulsé en rythme de l'ouverture et la fermeture des soupapes. L'admission crée deux types de bruit : bruit de "bouche" et bruit rayonné. Le bruit de bouche survient lors de l'ouverture et la fermeture des soupapes d'admission, quand les ondes acoustiques se propagent dans le système d'admission, constituée d'un répartiteur, un filtre à air et des tubulures. Les bruits de bouche se situent à l'entrée du conduit et peuvent être optimisés, en plaçant des résonateurs de Helmholtz ou des silencieux dans les tubulures du circuit [11]. Le bruit d'admission rayonné dépend des matériaux constituant le conduit d'admission.

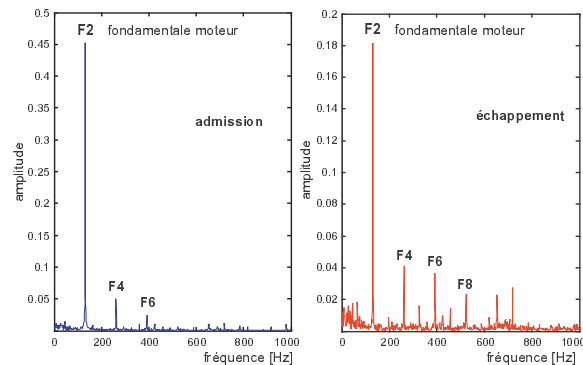


Fig. 2.12 – Admission/échappement à 4000 tr/min

Les harmoniques présentes pour un moteur 4 cylindres sont F2 et F4 (**Fig 2.12** et **Piste 7**), pour un 6 cylindres sont F1.5 F3 et F4.5. Le bruit d'admission peut être distingué en faisant varier l'ouverture

du papillon des gaz, le contact moteur coupé. Le bruit d'admission est audible également à l'extérieur du véhicule et il est soumis à des normes d'environnement.

2.2.4 Echappement

Le bruit d'échappement est produit par l'expulsion des gaz brûlés qui passent par les différentes parties de la ligne d'échappement comme le collecteur, le catalyseur et le silencieux. Cette expulsion a une fréquence fondamentale double de celle du régime moteur. A bas régime, les bruits de bouche comportent des harmoniques paires - F2, F4, F6, etc. (**Fig 2.12 et Piste 8**). A haut régime, la vitesse des gaz dans la ligne d'échappement crée des turbulences contribuant grandement dans les hautes fréquences du bruit.

2.2.5 Transmission

La transmission comprend la courroie, des galets et des poulies assurant la coordination entre les pièces rotatives du moteur (vilebrequin, arbres à cames, ...). La transmission est souvent associée à un bruit de sifflement hautes fréquences particulièrement audible à un régime élevé sur la face transmission du moteur. La courroie (ou la chaîne) assure l'entraînement de la distribution et génère des bruits de sifflement (ou des bruits dûs aux engrenages). Spectralement, ces bruits peuvent atteindre 2.5 kHz et au-delà. Le signal enregistré entre la courroie et le galet tendeur (**Piste 9**) montre que en haut régime les sifflements haute fréquence deviennent très importants (**Fig. 2.13**).

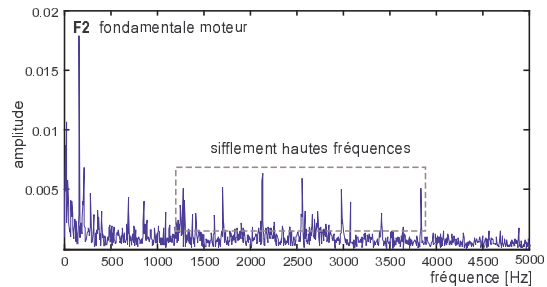


Fig. 2.13 – Transmission, à 5000 tr/min

2.3 Autres sources

Le moteur comporte également des organes auxiliaires qui sont aussi liés au cycle moteur. Il s'agit des accessoires moteur partiellement synchronisés avec la rotation du vilebrequin et contribuant au bruit moteur. Je note ici deux accessoires moteur possédant un bruit spécifique : l'alternateur et la pompe de direction assistée. Le troisième organe présenté, la boîte de vitesse, est un organe moteur à part.

2.3.1 Alternateur

L'alternateur a pour objectif de générer une tension continue afin d'alimenter en énergie électrique les différents éléments consommateurs, ainsi que la batterie. On distingue trois types de bruits générés par l'alternateur : sirènement, sifflement et bruit magnétique. Le sirènement est provoqué par l'écoulement d'air autour des pales des deux ventilateurs constituant l'alternateur. Le sifflement est émis par la courroie entraînant l'alternateur et synchronisée avec la vitesse de rotation du vilebrequin. Enfin, le bruit magnétique provient des vibrations des forces électromagnétiques entre les pôles de l'alternateur.

La vitesse de la rotation de l'alternateur est multiple (entier ou non entier) de la vitesse de rotation du vilebrequin. Ceci crée un intervalle plus ou moins dissonnant entre la fondamentale moteur et l'alternateur. La **Fig. 2.14** montre deux sons enregistrés à côté de l'alternateur d'un moteur à essence, respectivement à 4000 tr/min et 5000 tr/min. On peut constater dans ce cas que la fréquence fondamentale de l'alternateur suit l'harmonique F5 du moteur. Par ailleurs, le bruit de l'alternateur provoque une sensation de hauteur qui, pour l'exemple donné, se situe autour de 350 et 440 Hz. Ces fréquences correspondent aux sons Fa1 et La1. L'intervalle crée par les hauteurs de la fréquence fondamentale du moteur et de celle de l'alternateur est d'une octave et une quarte. Ces phénomènes peuvent être écoutés sur **Piste 11** et **Piste 12**.

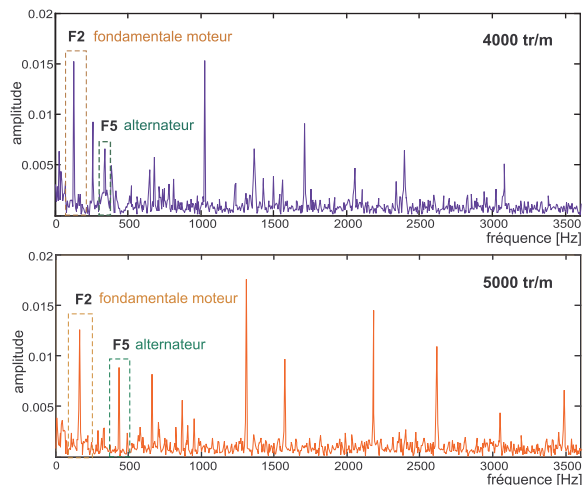


Fig. 2.14 – Bruit d'alternateur

2.3.2 Pompe de direction assistée

Tandis que l'alternateur assure l'alimentation en courant électrique, la pompe de direction assistée est chargée de maintenir le niveau de pression des liquides qui circulent dans le moteur. La pompe de direction assistée peut générer plusieurs types de bruits, dont les principaux sont caractérisés comme du grognement, du chuintement et du sirènement.

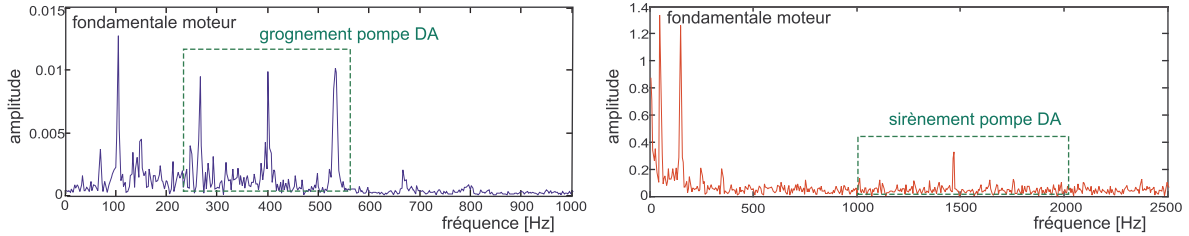


FIG. 2.15 – Bruit de grognement (à gauche) et bruit de sirènement (à droite) de la pompe DA

Le grognement est dû aux pulsations de la pression hydraulique et possède une fréquence fondamentale inférieure à 500 Hz (**Fig. 2.15** et **Piste 13**). Cette fréquence correspond à la rotation de la pompe et suit globalement le régime moteur. Le chuintement est un bruit large bande qui s'apparente à une chute d'eau. Le sirènement de la pompe DA est d'une fréquence plus élevée, et comporte des harmoniques dans la bande de fréquences de 1000 à 2000 Hz (**Fig. 2.15** et **Piste 14**).

2.3.3 Boîte de vitesse

La boîte de vitesse contrôle le rapport de couple. Son bruit est provoqué par l'engrenage des dents. On distingue le bruit de sirènement, lié à l'écart cinématique de l'engrènement, qui apparaît sous la forme de sons purs dans les moyennes et hautes fréquences (300-8000 Hz) (**Fig.3.2** et **Piste 10**). Le bruit de sirènement, dont les fréquences varient avec le régime moteur, dépend du rapport de boîte engagé.

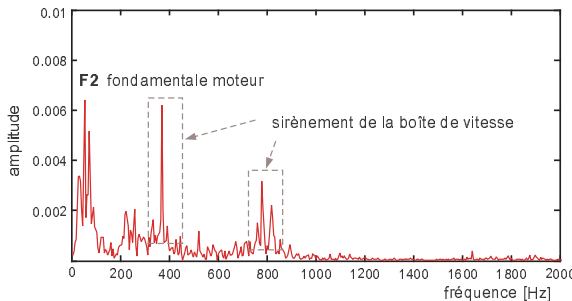


Fig. 2.16 – Sirènement de la boîte de vitesse

La boîte de vitesse produit encore un bruit dit de gaillonnement (en anglais "rattle"). Il trouve son origine dans le jeu existant entre les pignons, qui ne parviennent pas parfois à transmettre le couple moteur et provoquent un claquement métallique large bande. Ce bruit est fortement lié à l'irrégularité de rotation, l'acyclisme moteur. Il existe d'autres types de bruit de la boîte de vitesse, liés aux configurations, le passage par le point mort, etc. Globalement, la boîte de vitesse n'est pas une source provoquant un bruit de niveau sonore élevé. Toutefois, le cas contraire peut être observé.

Le bruit de la boîte de vitesse peut nuire à la qualité sonore dans l'habitacle, surtout pour les voitures haut de gamme, où le niveau des autres bruits est faible. Le bruit de la boîte de vitesse fait l'objet d'études acoustiques, qui se proposent de réunir les facteurs physiques et les paramètres perceptifs [12].

2.4 Identification des sources sonores

Cette description non exhaustive des sources acoustiques prouve la complexité sonore du bruit moteur. Pour distinguer la "signature" sonore de chaque source, il faut écouter le bruit à proximité de l'organe moteur émetteur. En pratique, les mécaniciens règlent "à l'oreille" le bon fonctionnement du moteur. L'intérêt de l'identification des sources sonores du moteur vient de l'objectif de maîtriser l'acoustique et d'apporter des modifications efficaces à un organe moteur précis. En règle générale, il est toujours préférable de traiter le bruit à la source.

Cette section parcourt les différentes méthodes utilisées dans la séparation des sources moteur. On peut distinguer deux types de méthodes : méthodes "physiques", modélisant le champ acoustique de la source et des méthodes "de signal", traitant le mélange sonore. Dans la première catégorie on peut citer l'holographie acoustique, dans la deuxième, l'analyse en composantes indépendantes (ICA).

Mesure des sources moteur La difficulté de la séparation des sources sonores vient de la proximité des organes moteur, ainsi que du fait que la majorité des bruits émis se recouvrent spectralement. Les mesures du bruit moteur se font soit dans les conditions de "banc moteur" dans une chambre anéchoïque, soit dans une situation réelle de conduite. La première solution donne l'avantage d'avoir des signaux dépourvus de bruits parasites et sans réflexion dans le compartiment moteur, alors que la deuxième offre plus de réalisme. Les mesures vibro-acoustiques, effectuées au début de ma thèse et qui ont fourni la majorité des exemples sonores dans ce manuscrit, sont décrites dans l'Appendice B.

Imagerie acoustique Les méthodes d'imagerie acoustique comme la mesure de pression, l'intensimétrie, la formation de voies ("beamforming") et l'holographie stationnaire ou non-stationnaire, estiment un champ acoustique autour d'une source sonore, en mesurant à l'aide d'un ou plusieurs capteurs, les paramètres physiques comme la pression, l'intensité ou la vitesse [13]. La méthode la plus utilisée dans l'industrie automobile pour la localisation des sources sonores est l'holographie. Son principe a été découvert par Gabor en 1948, et appliqué à l'acoustique en champ proche dans les années 1980 [14]. Il repose sur le calcul de la transformée de Fourier spatiale.

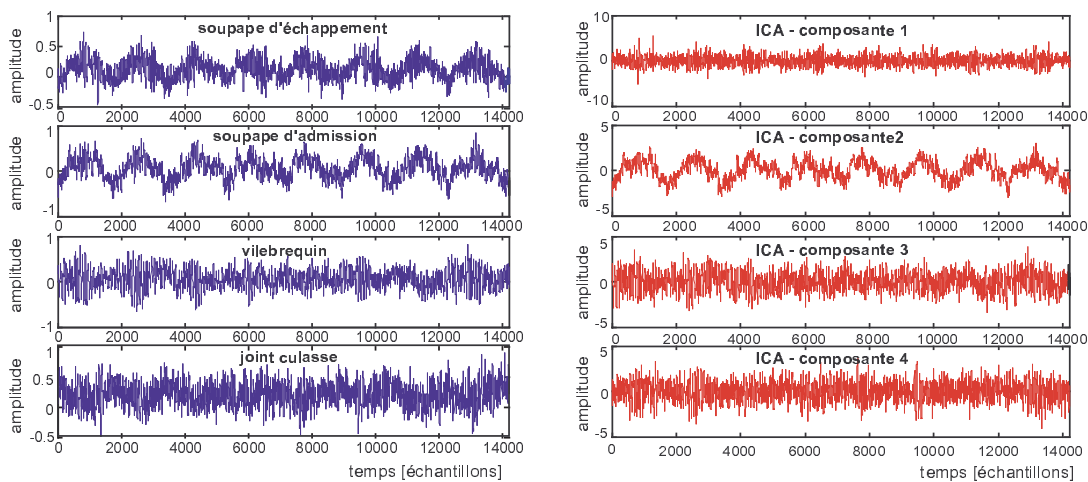


FIG. 2.17 – L'algorithme de séparation de signaux ICA appliqué sur 4 signaux moteur

Méthodes du signal Les méthodes du traitement du signal sont utilisées dans la séparation des sources à partir d'un mélange sonore. Ces méthodes trouvent essentiellement une application dans la musique et le traitement audio. On peut distinguer deux types de méthodes : méthodes "à l'aveugle" et méthodes avec un signal de référence. La méthode la plus élaborée de séparation de sources parmi les méthodes "à l'aveugle" est l'analyse en composantes indépendantes (ICA-"indépendant component analysis"), qui est basée sur l'analyse statistique des signaux [15]. La méthode de séparation de sources ICA trouve son application dans l'identification des sources sonores du moteur [16]. J'ai testé l'algorithme FastICA [17] sur 4 signaux enregistrés à proximité de différents organes moteur. Quatre composantes principales sont extraites, dont une de nature tonale (**Fig. 2.17**). Toutefois, on ne peut pas associer une source physique du moteur à chaque composante extraite.

En effet, la séparation "à l'aveugle" exige l'indépendance entre les sources sonores, une condition difficilement satisfaite à cause de la forte corrélation entre les parties moteur. Cette corrélation provient, comme on l'a déjà vu, du fait que les organes moteur obéissent à la même loi de rotation du vilebrequin et leurs émissions se recouvrent spectralement. Par ailleurs, pour effectuer une ICA proprement, le nombre de capteurs doit être au moins égal au nombre des sources potentielles. Or, il n'est pas possible de déterminer exactement le nombre de sources dans le cas du bruit moteur. En conclusion, l'identification et la séparation des sources moteur à l'aide des méthodes de traitement du signal reste un problème ouvert.

Résumé L'information la plus importante à retenir de cette présentation : le bruit moteur est le résultat du mélange de plusieurs sources sonores, qui possèdent chacune une signature sonore spécifique et dont les spectres se recouvrent souvent.

Conclusion sur l'acoustique moteur

Dans ce chapitre j'ai présenté les principaux mécanismes d'excitation à l'origine du bruit moteur. Les sources primaires sont les forces de pression et d'inertie, ainsi que les pressions pulsées. Elles provoquent des vibrations sollicitant les parties mobiles et fixes de la structure du moteur. L'acoustique du moteur est complétée par la contribution des sources secondaires qui assurent la distribution (arbres à cames, soupapes), l'injection (injecteurs), la transmission (courroie, galets), ou d'autres fonctions annexes du moteur (boîte de vitesse, alternateur). Chaque source moteur possède une signature sonore spécifique qui peut être perçue à proximité de l'organe émetteur. Les émissions acoustiques des organes moteur recouvrent souvent une plage fréquentielle similaire, ce qui rend difficile la séparation des contributions des différentes sources et confirme la nature très complexe du bruit moteur.

Orientation de la thèse : L'impossibilité de séparer complètement toutes les sources du bruit moteur m'a amenée à proposer une approche plus schématique du traitement du bruit moteur, basée sur le découpage temps-fréquence du cycle moteur et non pas sur la séparation spectrale des sources physiques. En effet, le positionnement temporel des événements dans le cycle moteur peut témoigner de la source sonore correspondante. Dans ce sens, le découpage schématique proposé plus loin garde la signification physique des événements dans le cycle moteur. Le concept détaillé est présenté dans la troisième partie de la thèse.

Acoustique véhicule

Le moteur participe grandement au bruit global du véhicule. Toutefois, il n'est pas la seule contribution sonore. L'acoustique de la voiture est déterminée aussi par deux autres contributions considérables : le bruit de roulement et le bruit aérodynamique (**Fig. 3.1**). Le bruit de roulement et le bruit aérodynamique sont importants, car à une vitesse élevée, leur apport prédomine sur le bruit du moteur seul. Pour une vitesse inférieure à 50 km/h, le groupe moto-propulseur est considéré comme la source majeure du bruit dans l'habitacle. Au-delà de cette vitesse, les bruits de roulement et le bruit aéro sont prépondérants. Pour des vitesses supérieures à 100-120 km/h, le bruit aérodynamique est la contribution la plus importante.

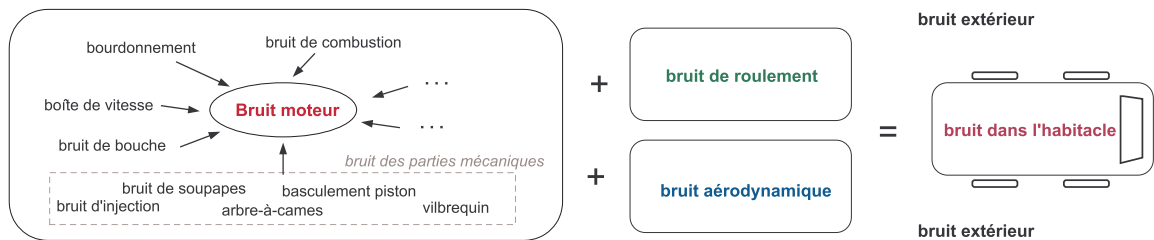


FIG. 3.1 – Schéma de l'acoustique véhicule : bruit moteur, bruit de roulement et bruit aéro

L'étude de l'acoustique véhicule s'intéresse à deux types d'émissions : à l'intérieur de l'habitacle et à l'extérieur de la voiture. L'acoustique dans l'habitacle est orientée vers la satisfaction du conducteur en termes de qualité sonore. L'émission sonore à l'extérieur du véhicule fait partie de l'étude environnementale sur les pollutions sonores en milieu urbain et est soumise à des normes gouvernementales [18].

3.1 Bruit de roulement

Le bruit de roulement, comme son nom l'indique, est provoqué par le contact des pneus et de la chaussée lors du déplacement du véhicule. Trois facteurs majeurs entrent en jeu dans la production du bruit de roulement : le pneu, la chaussée et la vitesse du véhicule.

Influence de la vitesse La vitesse est un facteur prépondérant, car il amplifie avec +3 dB(A) le niveau du bruit de roulement pour toute augmentation de 25 % de la vitesse du véhicule (**Fig. 3.2**) [19]. A partir de 50-60 km/h, le bruit de roulement est plus important que le bruit du moteur [20].

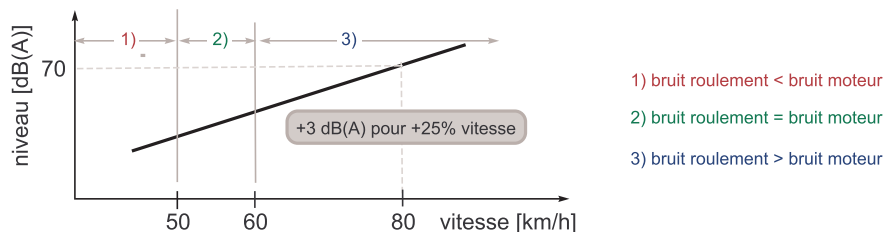


FIG. 3.2 – Relation entre le niveau du bruit de roulement et la vitesse du véhicule

Le niveau sonore du bruit de roulement des véhicules légers doit être compris, selon les normes, entre 69 et 75 dB(A) (le plus couramment 72 dB(A)), mesuré à 80 km/h à une distance de 7,5 m du véhicule. Une des mesures de la réduction du bruit de roulement consiste à limiter la vitesse afin de satisfaire les seuils de bruit de 60 dB(A) (le jour) et de 55 dB(A) (de nuit) en milieu urbain.

Structure du pneu Quatre facteurs influent les propriétés du pneu : la dimension, la construction, la composition et le dessin. Les recherches montrent que l'augmentation du diamètre et la diminution de la largeur du pneu contribuent à le rendre plus silencieux. Cependant, c'est la configuration inverse qui

est utilisée par les constructeurs, pour des raisons de performance. Côté type de construction, les pneus à carcasse radiale se révèlent plus silencieux de 0.7 dB(A) que ceux à carcasse diagonale. Les constructeurs de pneus évaluent le bruit rayonné à partir de l'étude du dessin du pneu sur sa bande de roulement. Toutefois, la qualité acoustique n'est pas prioritaire devant la sécurité et les performances du pneu.

Texture de la chaussée Le revêtement de la chaussée est un facteur important qui influence le bruit de roulement d'une manière plus importante que la structure des pneumatiques. Entre les pneus les plus silencieux et les plus bruyants, il n'existe qu'un écart inférieur à 4 dB(A), alors que cet écart peut être de l'ordre de 10-20 dB(A) pour des revêtements routiers différents. Les propriétés de la chaussée sont déterminées par trois facteurs principaux : la texture, l'absorption acoustique et la raideur du contact. L'absorption utilise la porosité de la chaussée pour atténuer le bruit.

Contact pneu/chaussée Les phénomènes générant le bruit de roulement sont de deux natures principales : l'une est purement vibratoire, l'autre est liée à la résonance d'air dans les cavités du pneu. Le bruit d'origine vibratoire se produit dans les basses et les moyennes fréquences (80 Hz-2 kHz) alors que le bruit d'origine aérodynamique se situe dans les moyennes et les hautes fréquences (1-5 kHz) [21]. Une classification plus précise distingue quatre mécanismes responsables du bruit de roulement : mécanique, aérodynamique, de friction et de propagation (**Fig.3.3**) [22]. Les phénomènes apparus lors du contact pneu/chaussée évoluent avec la vitesse du véhicule. Le spectre du bruit de roulement possède souvent un pic autour de 1 kHz, et garde généralement le même profil en fonction de la vitesse [23].

mécanisme	effet		
mécanique	vibrations radiales		vibrations tangentielles
friction	"stick-slip"		"stick-snap"
aérodynamique	turbulence d'air	résonance de tube	"air-pumping"
propagation	"horn effect"	absorption	directivité

FIG. 3.3 – Les principaux mécanismes à l'origine du bruit de roulement

3.1.1 Processus mécanique

Les vibrations radiales et transversales Le contact de la bande de roulement avec la surface rugueuse de la route crée deux types de vibrations : radiales et transversales (**Fig. 3.4**). Les vibrations radiales sont sollicitées suite aux impacts subis par la bande de roulement du pneu lors du contact avec la route. Le processus inverse, i.e la rupture de contact entre le pneu et la chaussée, provoque des glissements et des frottements de la surface, ce qui se traduit par des vibrations transversales de la bande de roulement [24]. Les vibrations transversales prennent la direction de la tangente du pneu dans le point sollicité et sont encore appelées vibrations tangentielles.

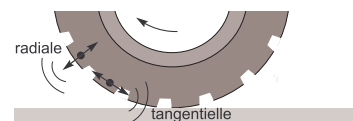


Fig. 3.4 – Vibration radiale et transversale

Les vibrations de la carcasse La bande de roulement interagit avec la surface de la route et transmet les excitations sur les côtés du pneu. Les études montrent 4 zones importantes dans l'émission de bruit (**Fig.3.5**). Deux de ces zones sont situées aux bords amont et aval de la bande de roulement et contribuent dans les fréquences (500 - 2 000 Hz). La troisième zone est disposée dans le passage de roue et génère un bruit de (500-600 Hz). Enfin, les côtés du pneu participent dans l'émission dans le spectre de 400 - 600 Hz et sont une source majeure de bruit. Le bruit de la vibration de la carcasse occupe principalement les basses et les moyennes fréquences, représentant environ 60-80 % de l'énergie totale du spectre [25].

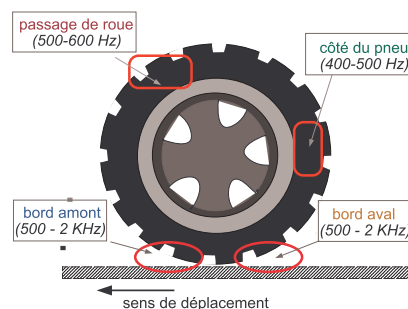


Fig. 3.5 – Vibrations de la carcasse du pneu

Les paramètres influents ici sont le dessin de la bande, la microtexture de la chaussée et le degré d'adhésion et de frottement pneu/chaussée. Les vibrations restent dans les basses fréquences (au-dessous de 1 kHz), car pour ce type de sollicitations, le pneu se comporte comme un filtre passe-bas.

3.1.2 Résonance d'air

Turbulence Lors du déplacement du véhicule, l'air s'écoule autour du pneu et crée des turbulences. Ce type de bruit, pas très significatif, a son origine dans le mouvement des pneus du véhicule, mais peut être mis dans la catégorie des bruits aérodynamiques. A une vitesse élevée et sur une route à surface poreuse, ce mécanisme peut contribuer dans les hautes fréquences [26].

"Air-pumping" Le phénomène de "air-pumping" se produit quand l'air, détenu dans les cavités du pneu n'entrant pas en contact avec la chaussée, est évacué suite à une compression lors du roulement (Fig. 3.6). Les études montrent que ce phénomène augmente grandement le niveau du bruit. En effet, la vitesse des jets d'air expulsés cause une grande pression. Le spectre du bruit rayonné est amplifié particulièrement dans la zone 1-3 kHz, et davantage sur une route non poreuse et relativement lisse.

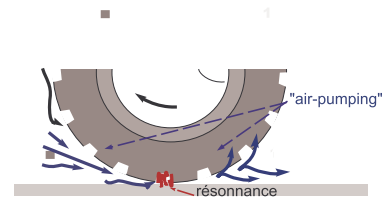


Fig. 3.6 – Le mécanisme d'"air-pumping"

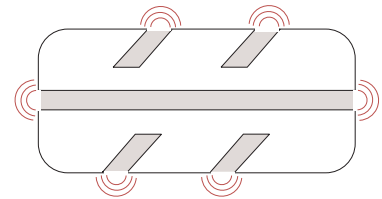


Fig. 3.7 – Résonance de Helmholtz

Résonances de Helmholtz Les résonances de Helmholtz peuvent apparaître au bord de la bande de roulement car les rainures du pneu forment des cavités, à l'extrémité desquelles l'air résonne (Fig. 3.7). Ceci provoque des rafales sonores au passage de chaque relief sur le sol. La résonance du tuyau est due aux ondes stationnaires qui se créent dans les rainures de la bande de roulement. Leurs fréquences de résonance dépendent du dessin du pneu, mais pas de la vitesse de rotation.

La longueur d'onde va être la double de la longueur du tuyau, si la rainure est ouverte dans les deux extrémités, et quadruple si l'ouverture est seulement dans un seul bout. Ce bruit contribue fortement au bruit de roulement total. Un phénomène similaire se produit entre les irrégularités de la route, mais ne semble pas avoir une conséquence sur le bruit.

Résonance dans le pneu Les résonances de la chambre à air à l'intérieur du pneu contribuent aussi dans le bruit généré. Ce comportement est perceptible lors du passage du pneu sur une route pourvue de discontinuités comme des rails ou des ponts, mais ne se manifeste pas pendant le roulement sur une route libre d'obstacles.

3.1.3 Propagation

"Horn effect" La surface de la route et l'arrière du pneu forment une cavité qui ressemble à un cor ("horn" en anglais). Ce phénomène amplifie de 20 dB le niveau sonore du bruit [27]. L'amplification la plus importante est autour de 2 kHz [28].

Directivité Le pneu n'a pas une forme idéalement sphérique et l'émission du bruit n'est pas la même dans toutes les directions. Globalement, les niveaux sonores en amont du pneu sont 1 dB plus élevés qu'en aval.

Absorption Selon les principes fondamentaux de la propagation en champ libre, la propagation du bruit s'atténue avec l'éloignement de la source, selon une loi qui dépend de la forme du front d'onde. Dans le cas idéal, la pression acoustique décroît selon la loi inverse quadratique en fonction de la distance. En réalité, le bruit de roulement est également réfléchi par le sol. Dans ce processus entre en compte la structure de la route en termes d'épaisseur du revêtement et de porosité. Sur une surface poreuse une partie du bruit propagé est absorbée, au sens où l'énergie acoustique est transformée en énergie thermique.

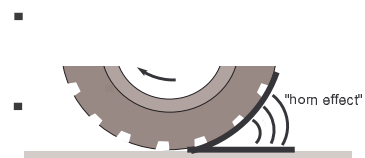


Fig. 3.8 – Le mécanisme de "horn effect"

3.1.4 Friction

Stick-slip Le phénomène de stick-slip (ou encore de glissement) provient des forces de friction qui se créent lors du contact entre la bande de roulement et la surface de la chaussée. La friction entre le pneu et la route est une combinaison de deux composantes : l'adhésion et l'hystérésis. L'adhésion détermine la tendance du pneu à coller à la surface et dépend de la composition du matériau du pneu et de la rugosité du revêtement de la chaussée. L'hystérésis produit un effet inverse. Dans ce cycle d'adhérence et de rupture d'adhérence se crée un effet de ventouse provoquant du bruit.

Stick-snap Cet événement se produit quand la route devient glissante. La perte d'adhésion cause des vibrations et probablement, un flux d'air entre le pneu et la route. Le bruit résultant n'est pas considérable, par rapport aux autres mécanismes.

3.1.5 Exemple de bruit de roulement

Les particularités sonores du bruit de roulement en fonction de 3 paramètres (vitesse, chaussée, position) sont exposées sur la **Fig. 3.9**. Les signaux ont été enregistrés à deux vitesses : en 3-ème à 50 km/h et en 4-ème à 90 km/h, sur deux pistes : piste "standard" et piste "gros grain" et dans deux positions : à l'intérieur dans l'habitacle et à l'extérieur près de la roue (**Piste 15-22**). Le niveau sonore dans l'habitacle est 20-30 dB moins élevé que le niveau à l'extérieur du véhicule. Les niveaux sont aussi plus élevés à une vitesse supérieure. Il semble sur cet exemple qu'avec l'augmentation de la vitesse, la différence entre les revêtements est moins exprimée. Ainsi, à 50 km/h, le niveau de la piste "gros grain" est supérieur au niveau de la piste "standard", alors que ceci n'est plus le cas à 90 km/h. En conclusion, le bruit de roulement dépend des interactions entre la chaussée et le pneu. Il croît avec la vitesse et diffère selon la position d'émission (intérieur/extérieur).

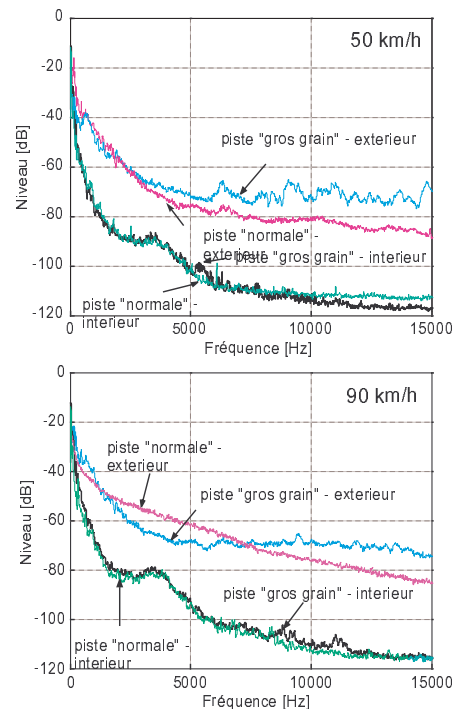


Fig. 3.9 – Bruit de roulement

3.2 Bruit aérodynamique

Le bruit aérodynamique est le résultat des écoulements d'air autour de la voiture. Ce bruit dépend fortement de la vitesse et la forme du véhicule et devient prépondérant à plus de 100 km/h. Dans l'habitacle, à une vitesse élevée, le bruit aérodynamique prédomine le bruit moteur et le bruit de roulement. D'un point de vue spectral, le bruit aérodynamique est un bruit large bande, généralement situé au-dessus de 400 Hz et contribuant dans le spectre dans les moyennes et surtout les hautes fréquences [29].

3.2.1 Principaux mécanismes

Le bruit aérodynamique est le résultat de deux mécanismes principaux : aéroélastique et aéroacoustique (**Fig.3.10**). Le bruit d'origine aéroélastique se situe fréquemment plus bas, alors que le bruit aéroacoustique est plus marqué dans les moyennes et les hautes fréquences, dans la bande [300-1000 Hz].

Efforts aéroélastiques Le premier mécanisme exprime le rayonnement de la structure du véhicule provoqué par les fluctuations de l'écoulement autour du véhicule. L'air vient exciter la structure de la voiture. Les vibrations ainsi engendrées se propagent à l'extérieur, mais également à l'intérieur dans l'habitacle. Ce processus vibratoire prend en compte plusieurs paramètres comme les propriétés mécaniques des matériaux et les lois de couplage fluide-structure.

Sources aéroacoustiques Le deuxième mécanisme concerne le rayonnement acoustique provoqué par l'écoulement turbulent. L'avancée de la voiture dans l'air et la résistance qu'elle subit lors de ce mouvement crée des zones de turbulence. De plus, puisque le véhicule ne présente pas un profil parfaitement lisse, les parties dépassant ce profil (rétroviseurs, essuie-glaces, joints) sont susceptibles de créer de fortes perturbations turbulentes.

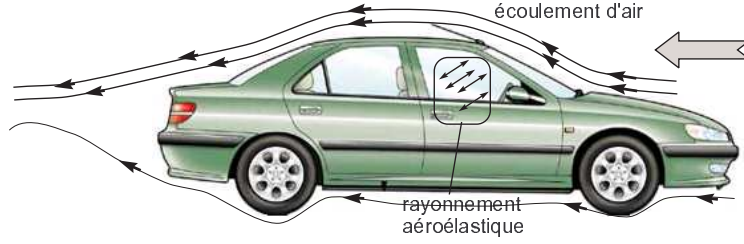


FIG. 3.10 – Mécanismes de sollicitation aéroélastique et aéroacoustique

3.2.2 Modélisation physique

Les études sur le bruit aérodynamique prennent en compte une modélisation théorique des processus, à l'aide des principales équations de l'hydrodynamique. Cette théorie donne une bonne base pour poursuivre les recherches avec d'autres approches plus expérimentales. La modélisation physique est orientée vers les deux mécanismes principaux : les efforts aéroélastiques que subissent les parois du véhicule et les phénomènes aéroacoustiques, créés lors de l'écoulement d'air.

Efforts aéroélastiques Les fluctuations de pression pariétale génèrent des efforts mécaniques sur les panneaux de la carrosserie de la voiture. Les fluctuations de pression sont essentiellement contrôlées au voisinage de l'écoulement (pour des nombres de Mach faibles) et s'expriment avec l'équation de Poisson

$$\Delta^2 p = \gamma \quad (3.1)$$

où le terme source γ est une fonction connue de la vitesse instantanée. Ces pressions ne contribuent pas au champ acoustique lointain, car leur amplitude décroît rapidement en s'éloignant de l'écoulement. La modélisation physique des efforts aéroélastiques passe par une représentation aléatoire du chargement. Ainsi, l'excitation due aux fluctuations de pression pariétale,

$$S_{pp}(x, \xi, f) = \Phi_{pp}(x, f) \cdot \gamma(x, \xi, f) \cdot e^{-\frac{i2\pi f\xi}{U_p(x, \xi, f)}}, \quad (3.2)$$

est modélisée par des fonctions aléatoires de l'espace et des corrélations spatiales [29]. Les premières sont caractérisées par leur amplitudes, les deuxièmes, dépendent de la fréquence f . Dans la relation 3.2, $\Phi_{pp}(x, f)$ est l'autospectre des fluctuations de pression au point x , $\gamma(x, \xi, f)$ est la fonction de cohérence entre deux points séparés de ξ et $U_p(x, \xi, f)$ est la vitesse de convection dans la direction nominale de l'écoulement. La forme de ces trois fonctions est très importante pour le calcul de la réponse vibro-acoustique des parois.

Sources aéroacoustiques La génération du bruit par l'écoulement traduit le transfert qui s'effectue entre deux types d'énergie - cinétique et acoustique. La modélisation théorique du bruit aéroacoustique trouve ces bases dans les travaux de Lighthill [30]. Il ajoute à l'équation de propagation d'onde des termes sources et considère les fluctuations dans un champ acoustique [31]

$$\frac{\partial^2 \rho'}{\partial t^2} - c_0^2 \frac{\partial^2 \rho'}{\partial x_i^2} = \frac{\partial q}{\partial t} - \frac{\partial F_i}{\partial x_i} + \frac{\partial^2 T_{ij}}{\partial x_i \partial x_j}. \quad (3.3)$$

Ici c_0 est la vitesse du son, ρ' est la masse volumique, le terme $\frac{\partial q}{\partial t}$ décrit le bruit émis par le déplacement de volume dû au déplacement de surface rigide, alors que le terme $\frac{\partial F_i}{\partial x_i}$ représente les sources dues aux forces aérodynamiques exercées par la surface solide. Le tenseur de Lighthill, $\frac{\partial^2 T_{ij}}{\partial x_i \partial x_j}$, représente les sources de bruit ayant pour origine les effets de turbulence et de viscosité du milieu. Sans être obligé de résoudre l'équation, il est possible d'estimer l'ordre des grandeurs. Ainsi, on obtient une proportionnalité entre le module de l'intensité acoustique et la puissance 8 de la vitesse [5]. L'équation de Lighthill est valable dans la considération d'un champ lointain, i.e quand la source n'est pas proche des parois. Il est également nécessaire de tenir compte des réflexions d'onde acoustique sur les parois. Une extension du modèle de Lighthill est donnée par Curle, qui inclut les réflexions aux bords [32]. Plus tard, Powell tient compte des parois solides et ajoute les fluctuations turbulentes se produisant à proximité des surfaces [33]. En résumé, la modélisation du bruit aérodynamique est un problème très complexe.

Exemple de bruit aéroacoustique Le spectre du bruit aéroacoustique occupe une bande fréquentielle relativement large (**Fig. 3.11**). Ces exemples présentent la différence de spectre avec et sans la contribution du bruit aéroacoustique (**Fig.3.11 a**) et le rapport entre le bruit aéroacoustique isolé et le bruit total de la voiture (**Fig.3.11 b**).

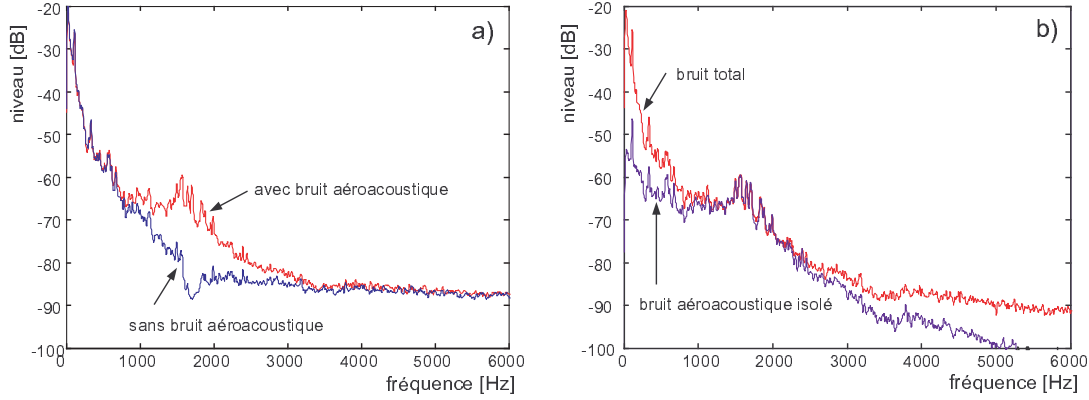


FIG. 3.11 – Spectre du bruit aéroacoustique par rapport au bruit total

On peut observer une présence du bruit aéroacoustique dans la plage 1-3 kHz, mais également une contribution dans les plus basses fréquences. L'exemple sonore associé à ces graphiques et qui compare les paires "bruit total - bruit aéro isolé" et "bruit total - sans bruit aéro", se trouve sur la **Piste 23**.

3.2.3 Prédiction du bruit aérodynamique

L'étude et la conception aérodynamique des véhicules ont trois cibles : le bruit de forme, dû aux écoulements autour de la voiture, le bruit provoqué par les accessoires (rétroviseur, antenne, etc.) et le bruit d'étanchéités et de joints. L'objectif de l'aérodynamique est de minimiser la puissance P que le véhicule perd pour vaincre la résistance à l'avancement dans l'air, notée

$$P = \frac{1}{2} \rho S V^3 C_x. \quad (3.4)$$

Cette formule montre que la puissance perdue dépend, hormis de la masse volumique de l'air ρ , de la surface frontale du véhicule S , du cube de la vitesse de déplacement V^3 , et du coefficient de traînée, C_x , appelé encore coefficient de pénétration dans l'air. Les voitures actuelles possèdent un coefficient C_x de l'ordre de 0.3 à 0.4 [34].

Profil géométrique L'étude du bruit aérodynamique se fait dans de grandes souffleries reproduisant d'une manière très fidèle l'avancement de la voiture dans l'air. Ces conditions expérimentales sont utilisées pour tester des maquettes, comportant la forme géométrique quasi-finale de la voiture. D'autres essais plus théoriques se font dans l'objectif de recueillir des données sur les caractéristiques aérodynamique de formes précises, dont les paramètres géométriques sont contrôlés. La différence entre le bruit aérodynamique généré par les profils de trois voitures différentes (**Fig. 3.12** et **Piste 24**) est bien exprimée dans la zone 1-4 kHz.

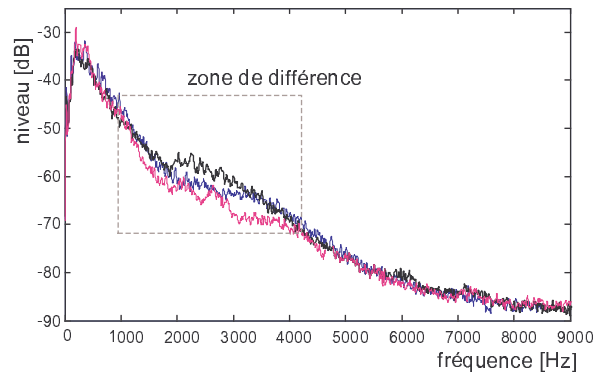


Fig. 3.12 – Bruit d'écoulement de 3 voitures

3.3 Chemins de transfert

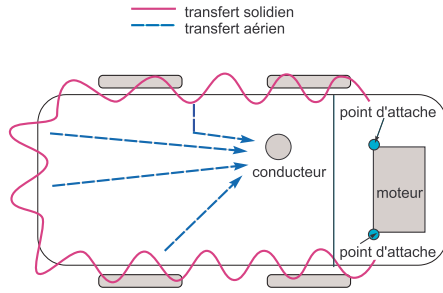


Fig. 3.13 – Chemins de transfert du bruit moteur

Les excitations en provenance des trois contributions du bruit automobile, le bruit moteur, le bruit de roulement et le bruit aérodynamique, se propagent via deux chemins de transfert : solidien et aérien. Cet acheminement est représenté par l'équation standard de la propagation des ondes acoustiques, avec deux constantes de la vitesse de propagation : 5000 m/s dans l'acier pour la voie solide et 340 m/s pour l'air pour la voie aérienne. Une illustration de la propagation solidienne et aérienne est donnée sur la (Fig. 3.13)[35]

3.3.1 Transfert solidien

Le chemin de transfert solidien se fait à travers la structure de la voiture, ce qui explique son autre appellation, bruit de structure. Les excitations générant ce type de bruit sont en provenance de deux sources : le moteur et le roulement.

Transfert du bruit moteur La transmission vers l'habitacle du bruit émis dans le compartiment moteur passe par les points d'attache (on encore la suspension) du moteur. D'un côté, les points d'attache doivent assurer la stabilité de fixation du moteur, alors que d'un autre côté, ils doivent également permettre de petites oscillations du moteur autour de ses trois axes. Les vibrations au niveau des cales moteur sont transmises à la carrosserie de la voiture, qui par la suite crée le rayonnement acoustique sous la forme d'un champ d'ondes stationnaires dans l'habitacle. Les études d'isolation contre le bruit solidien concernent essentiellement les modes de vibrations au-dessous de 200-400 Hz.

Exemple 1 Les excitations vibratoires de la voiture dues au moteur sont exposées sur la Fig. 3.14. Deux situations sont étudiées : avec le moteur en marche et le moteur coupé. Les mesures sont effectuées sur une piste lisse afin de minimiser les vibrations dues au roulement. Trois points représentatifs témoignent de la transmission des vibrations : la cale moteur, la glissière du siège conducteur et le plancher dans l'habitacle derrière le conducteur (voir Appendice B). En absence du moteur, on peut observer une excitation de la structure nettement moins importante. Il est incontestable, que la cale moteur est soumise à l'excitation la plus importante de la part du moteur, alors que les vibrations du siège et du plancher possèdent des modes d'excitation spécifiques dues au moteur. Ces régions sont situées au-dessous de 400 Hz pour le plancher et 200-400 et 400-600 Hz pour le siège conducteur. Les signaux correspondant aux graphiques (Fig. 3.14 a, a', Fig. 3.14 b, b' et Fig. 3.14 c, c') se trouvent respectivement sur Piste 25, Piste 26 et Piste 27.

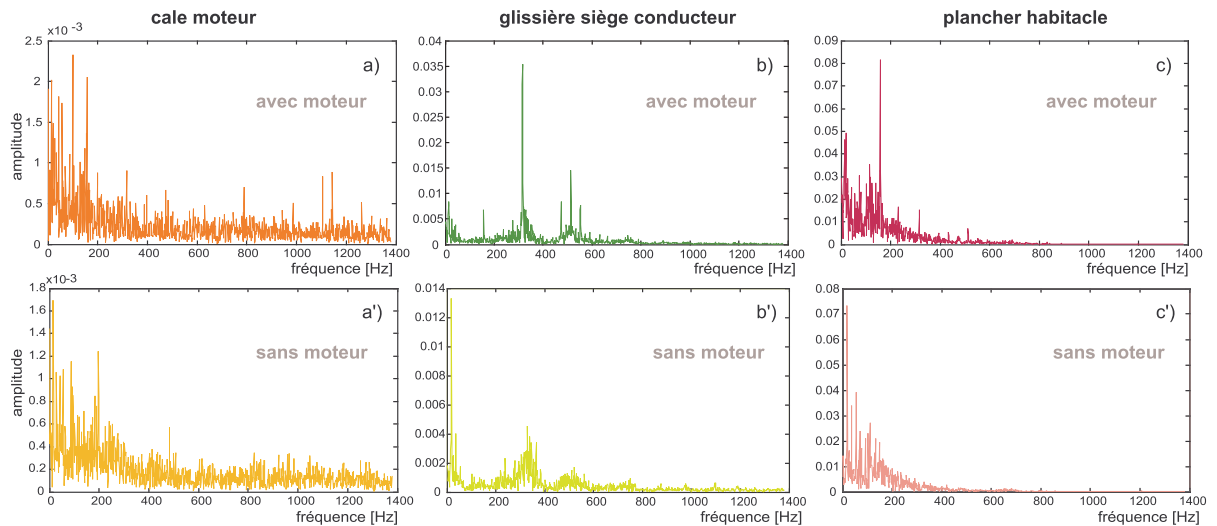


FIG. 3.14 – Excitations vibratoires de la voiture avec et sans moteur

Transfert du bruit de roulement La deuxième importante contribution dans le transfert solide vient du roulement. Les excitations de la route sont transmises à la structure de la voiture à travers le système de suspension. Ce processus s'identifie à des vibrations dans le plancher et des sièges du véhicule. Une situation particulière étudiée lors du franchissement des obstacles ou des irrégularités sur la route est le bruit de percussion. Il faut mentionner que le moteur joue un rôle d'absorbeur dynamique des vibrations dues aux irrégularités de la route dans la gamme de fréquences 4-8 Hz. Par ailleurs, le corps humain est très sensible dans ce domaine fréquentiel. Enfin, les excitations d'origine aéroélastique viennent s'ajouter aux sollicitations de la structure de la voiture.

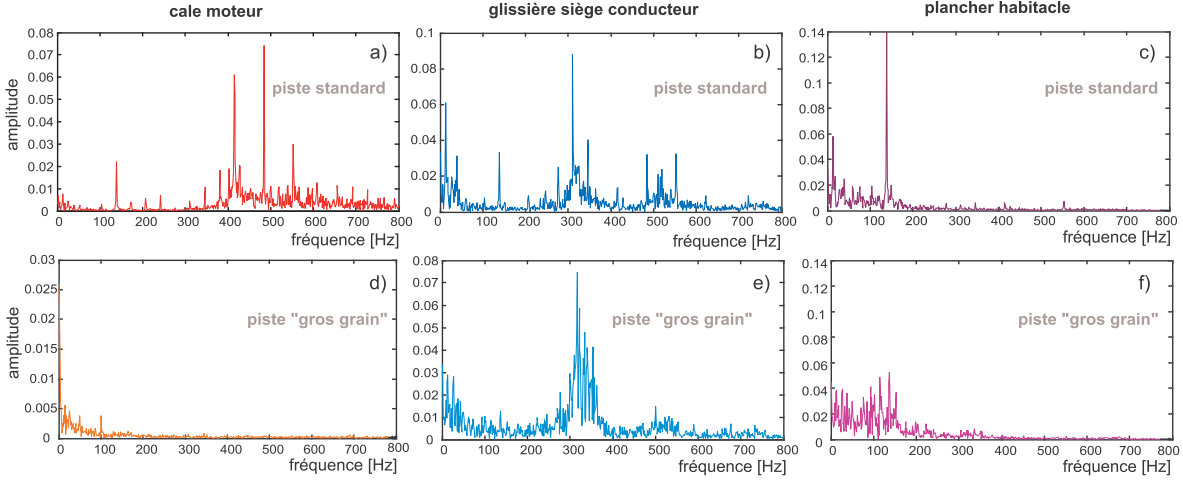


FIG. 3.15 – Excitation vibratoire sur piste "normale" et sur piste "gros grain"

Exemple 2 Cet exemple de transfert solide concerne la transmission vibratoire en trois points (cale moteur, glissière, plancher) lors du roulement de la voiture sur deux pistes différentes : une piste à revêtement standard et une piste à revêtement "gros grain", à une vitesse de roulement de 90 km/h et à un rapport de vitesse 3 (Fig. 3.15). Les sollicitations dues aux irrégularités de la route affectent surtout le plancher de l'habitacle. En revanche, les vibrations sur la cale moteur deviennent plus faibles, probablement grâce à la compensation du moteur par rapport aux irrégularités de la route. Les signaux correspondant aux graphiques de (Fig. 3.15 a-e) se trouvent respectivement sur **Piste 28-33**.

3.3.2 Transfert aérien

Le deuxième chemin de propagation des bruits automobile est la voie aérienne. C'est l'air qui porte les bruits et vibrations, émis dans le compartiment moteur. Le bruit aérodynamique est transmis via les parois de la voiture dans l'habitacle. Ce transfert dépend de l'étanchéité des joints des portières, vitres, etc.

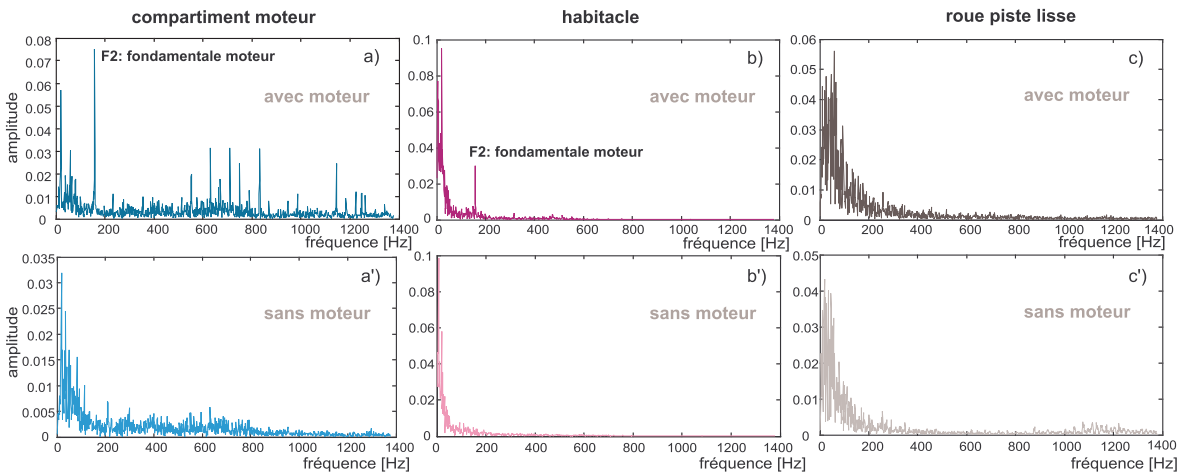


FIG. 3.16 – Transfert aérien entre le compartiment moteur, le roulement et l'habitacle

Exemple 3 L'exemple de la **Fig. 3.16** donne une idée du transfert aérien entre le compartiment moteur et l'habitacle, ainsi que du bruit de roulement extérieur, en présence ou en absence du moteur. Les conditions d'enregistrement sont identiques à celles de l'Exemple 1. Les signaux correspondant aux graphiques **Fig. 3.16 (a, a')**, **Fig. 3.16(b, b')** et **Fig. 3.16 (c, c')** se trouvent respectivement sur **Piste 34**, **Piste 35** et **Piste 36**.

3.4 Ambiance sonore dans l'habitacle

Les trois contributions principales du bruit automobile, en provenance du moteur, du roulement et d'écoulement d'air, représentent la quasi-totalité de la problématique acoustique véhicule. Toutefois, parmi les autres contributions sonores on peut compter plusieurs bruits et sons spécifiques.

Bruits et sons divers Un bruit particulier est le bruit de démarrage qui donne la première impression auditive du moteur. La ventilation et la climatisation provoquent des bruits aéroacoustiques à l'intérieur de la voiture. Ces bruits peuvent influencer la perception thermique dans l'habitacle. L'ambiance sonore dans l'habitacle est complétée avec les sons de fonctions et d'information (avertisseur sonore, clignotants, etc.) qui font également l'objet du design sonore. Parmi ces bruits "accessoires", les bruits dits ouvrants (ouverture et fermeture de porte, de toit, de vitre, etc.) ont un grand impact sur l'appréciation de la voiture. Parmi les autres accessoires sonores on peut citer le bruit d'essuie-glaces, les bruits de freins. Tous ces sons et bruits créent l'ambiance sonore générale de la voiture et contribuent à la qualité sonore.

Conclusion sur l'acoustique automobile

Le bruit global automobile est le résultat de trois contributions majeures : le bruit du groupe moto-propulseur, le bruit de roulement et le bruit provoqué par l'écoulement d'air. Plusieurs mécanismes très complexes sont à l'origine de ces bruits. L'impact des trois contributions est variable en fonction de la vitesse de la voiture et de l'état de la route. A une vitesse peu élevée, c'est le bruit moteur qui est prépondérant. En revanche, à partir de 50 km/h, le bruit de roulement commence à participer grandement, alors qu'à une vitesse très élevée (>100 km/h), le bruit aérodynamique est la source majeure. Les émissions vibratoire et acoustique des sources du bruit automobile sont transmises via deux chemins de transfert : solidien et aérien. Le chemin solidien est responsable du transfert dans les basses fréquences, via la structure du moteur, alors que la voie aérienne transmet plutôt les moyennes et les hautes fréquences. L'ambiance sonore dans l'habitacle de la voiture est un mélange des trois contributions, mais également d'autres bruits accessoires, relatifs aux fonctions d'avertissement du conducteur. L'étude de l'acoustique automobile prend en compte également le bruit émis à l'extérieur de la voiture afin de respecter les normes d'environnement en termes de pollution sonore. Enfin, la qualité sonore du bruit véhicule a pour objectif le confort acoustique du conducteur dans l'habitacle.

Orientation de la thèse : Le bruit du groupe motopropulseur est la source de bruit qui est exploitée d'une manière prioritaire dans la thèse. J'ai choisi de travailler en particulier sur les bruits émis directement dans le compartiment moteur, et pas dans l'habitacle, car mon objectif et celui d'un projet de recherche de mon entreprise était de se concentrer sur l'émission "brute" du bruit au banc moteur, sans se préoccuper de l'acheminement du bruit vers l'habitacle. En effet, l'étude de l'émission du moteur seul permet de comprendre les problèmes perceptifs des bruits moteur dès leur origine. Ainsi, la compréhension du bruit moteur au banc pourrait servir comme aide dans la phase préliminaire de la conception d'un moteur.

Qualité de bruit automobile

La qualité sonore fait l'objet de plusieurs études dans le domaine de la perception auditive. Cet intérêt, en vue des enjeux industriels, est particulièrement important dans le domaine de l'automobile. Ce chapitre présente la qualité sonore dans la continuité de la perception auditive des sons, mais aussi dans le contexte des bruits automobile. La présentation inclut les outils les plus utilisés dans l'étude perceptive des sons comme certains indicateurs psychoacoustiques, ainsi que la méthodologie des tests perceptifs.

4.1 Perception et qualité sonore

La qualité sonore représente une branche de la perception sonore orientée vers la recherche d'indices perceptifs qui assurent l'agrément sonore chez l'auditeur. Dans cette section vont être présentées les différentes approches de la perception, ainsi que les aspects de la qualité sonore dans le contexte automobile.

4.1.1 Démarche perceptive

Le domaine de la perception sonore comprend trois approches complémentaires : la psychoacoustique, la psychologie cognitive et la psychologie écologique (**Fig. 4.1**).

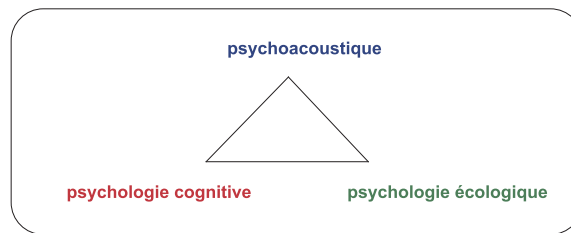


FIG. 4.1 – Les trois approches complémentaires de la perception sonore

Approche psychoacoustique La psychoacoustique étudie la relation entre les propriétés physiques des sons et les sensations auditives chez l'homme. Elle s'intéresse à l'anatomie et la physiologie de l'oreille et aux mécanismes de codage des sons dans la cochlée. Les études psychoacoustiques ont permis de déterminer des indicateurs psychoacoustiques, comme la sonie, la rugosité, l'acuité, etc., et d'établir des modèles, permettant de prédire la réaction du système auditif aux différents stimuli [36]. Dans un autre plan, la psychoacoustique s'intéresse aux processus de flux auditifs, i.e au groupement et à la ségrégation de signaux concurrents, émis simultanément ou en séquence [37]. Par ailleurs, l'existence des flux auditifs dans la perception des bruits moteur au ralenti sera le point de départ de notre étude perceptive, présentée dans la troisième partie de la thèse. Le traitement des flux auditifs est lié à la théorie des scènes auditives et aux processus de mémoire et de représentation mentale, un des piliers de l'approche cognitive [38].

Approche cognitive La psychologie cognitive étudie la manière dont le cerveau perçoit, mémorise et organise l'information. L'approche cognitive regroupe divers processus mentaux liés à la mémoire, le langage, le raisonnement, l'attention. La cognition est fondée sur le concept de la "représentation mentale", qui effectue le lien entre la réalité physique et le traitement des sons par le cerveau [39]. L'approche cognitive utilise le langage pour la description des attributs sonores et met en place des méthodes sémantiques (verbales) dans l'étude de la perception auditive. Cet outil est à la base du test de description verbale, présenté plus loin dans la section consacrée à la méthodologie des tests perceptifs.

Approche écologique Le concept de l'approche écologique, introduite par Gibson dans les années 1960-70, s'oppose à l'approche cognitive, en abandonnant la représentation mentale en profit de la perception directe de l'environnement [40]. L'approche écologique suppose l'individu lié à son environnement dans le sens où l'individu peut percevoir différemment les sons en fonction du contexte environnant. L'étude de l'approche écologique est concentrée sur la façon dont les structures émettrices des sons reflètent l'environnement. Cette supposition amène à établir des relations non seulement entre les sons et les

sensations perçues, comme le fait la psychoacoustique, mais entre les propriétés mécaniques de la source et les sensations perçues, comme le fait la nouvelle branche dans le domaine de la perception sonore, la psychomécanique [41].

4.1.2 Objectifs de la qualité sonore

La qualité sonore est une exigence relevée par l'industrie dans la recherche permanente d'amélioration d'un produit en vue de la satisfaction du client. La qualité sonore est un aspect de la qualité globale, relatif aux propriétés auditives des sons issus du fonctionnement efficace du produit. D'autres aspects de la qualité sensorielle globale sont la qualité tactile, visuelle ou olfactive. Les études de qualité sonore concernent des sons de nature différente comme les sons des instruments de musique, de divers bruits domestiques [42], ainsi que les bruits automobile [43].

Facteurs influants sur la qualité sonore La qualité sonore est un processus complexe où plusieurs facteurs entrent en jeu. D'une manière générale, on distingue deux types de facteurs : auditifs et émotionnels. Les facteurs auditifs sont liés à la perception des sons par le système auditif. Ces facteurs, exprimés à l'aide des indicateurs psychoacoustiques comme la sonie, la rugosité, l'acuité, etc., permettent de faire un lien entre la réalité physique des sons et la perception subjective (Fig. 4.2).

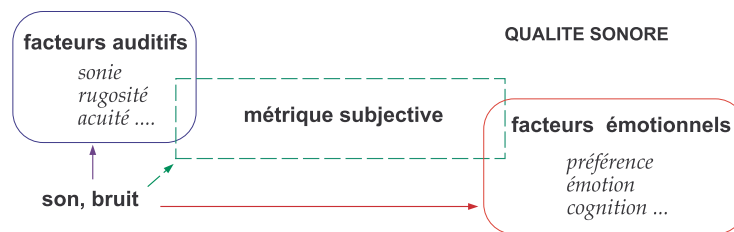


FIG. 4.2 – La qualité sonore subjective en fonction des facteurs auditifs et des facteurs émotionnels

Parallèlement, les facteurs dits émotionnels, ont un grand impact sur l'appréciation des sons. Ils décrivent l'état émotionnel (bonne humeur, fatigue, etc) et les préférences subjectives des auditeurs. Les facteurs émotionnels reflètent l'expérience psychologique, sociale et culturelle des auditeurs. La qualité sonore est donc une combinaison subtile entre ces deux types de facteurs. Leur interaction va créer une métrique subjective propre à chaque personne.

Objectif de la qualité sonore Le grand objectif de la qualité sonore est de rendre agréable, d'un point de vue sonore, l'utilisation du produit par le client. Concrètement, il est nécessaire de comprendre l'interaction entre les facteurs auditifs et émotionnels et de trouver les points communs entre les métriques subjectives de plusieurs personnes. La ressemblance entre les métriques subjectives permet de regrouper les sensations auditives d'une catégorie de gens, en créant des profils sonores. Notre étude ne portera pas sur les facteurs émotionnels, mais mettra l'accent sur l'étude des facteurs auditifs.

Facteurs auditifs Les facteurs auditifs, exprimés via les indicateurs psychoacoustiques, sont extraits à l'aide de tests perceptifs. Chaque indicateur décrit un aspect du timbre des sons, par exemple la sonie indique le niveau perçu et la hauteur exprime la perception du "grave-aigu". Les tests sur l'audition ont permis de déduire des expressions explicites pour certains indicateurs. Toutefois, ces formules sont obtenues souvent à l'aide de sons théoriques (sinusoïde, bruit blanc, sons harmoniques) et n'arrivent pas à expliquer complètement les propriétés des sons réels. Il est nécessaire donc de traiter à part chaque classe de sons afin d'extraire ses propres dimensions perceptives. Cette extraction se fait à l'aide de tests perceptifs.

4.1.3 Aspects de la qualité sonore automobile

La qualité sonore d'une voiture est liée d'une manière intrinsèque à la qualité globale. Très souvent, la qualité sonore est interprétée plutôt en termes de réduction de bruit. Or, il s'avère qu'elle représente un problème beaucoup plus complexe, car même si le niveau sonore du bruit reste le facteur le plus influant sur la gêne auditive, il n'est pas le seul responsable de la qualité sonore.

Automobile et identité sonore Les constructeurs automobile conçoivent une voiture en gardant une certaine cohérence entre la gamme, le design, l'utilisation de la voiture, ainsi que le type de clients pour lequel elle est conçue. Dans ce sens, la qualité sonore doit être en accord avec cette image globale de la voiture. En effet, le bruit représente un "passeport sonore" qui permet d'identifier la marque et le modèle de la voiture. Une voiture haut de gamme avec un bruit médiocre n'est pas bien perçue par le client, car ce point faible met en cause la qualité globale du véhicule. Pour les voitures bas de gamme, par souci d'un coût de fabrication moins important, les exigences sonores sont orientées juste vers un niveau sonore acceptable, plutôt que vers une acoustique particulièrement soignée.

"Bruit de qualité - bonne qualité" La bonne qualité sonore contribue à l'image positive de la voiture. En situation statique, le bruit de fermeture de porte est un exemple de test de la qualité d'une voiture. Un bruit de qualité permet d'imaginer la solidité des matériaux et peut créer un sentiment positif de sécurité chez le client (**Fig. 4.3**). D'une manière générale, un bruit grave, feutré et "énergique" donne une meilleure impression qu'un bruit aigu, claquant et "mou". Les bruits de porte sont un bon exemple de sons où le niveau sonore élevé non seulement ne nuit pas à la qualité, mais peut contribuer à l'appréciation positive du bruit

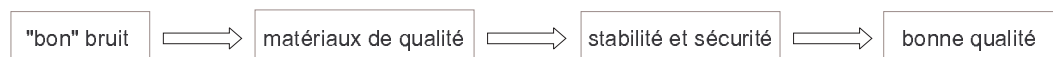


FIG. 4.3 – L'acheminement entre la perception auditive et la notion de qualité

En situation dynamique, c'est le bruit moteur qui est grandement responsable de la qualité sonore de la voiture. Le bruit moteur doit d'abord refléter le bon fonctionnement du moteur. Il me semble que la préoccupation du client pour le "bon fonctionnement" reste un des facteurs émotionnels les plus importants à prendre en compte. Ainsi, un bruit peut être mieux toléré s'il est associé au fonctionnement jugé "normal" du moteur. En revanche, même un bruit d'un niveau sonore relativement faible peut s'avérer très désagréable, s'il laisse supposer un dysfonctionnement de la voiture. Des bruits parasites, comme par exemple le bruit provoqué par un joint défectueux, un sifflement d'air à travers la portière, etc. sont considérés comme négatifs et nuisent à la qualité. Ce désagrément peut devenir d'autant plus important que la source d'émission de bruit reste non identifiable par l'oreille.

Niveau sonore et confort L'espace de l'habitacle est l'endroit ciblé par les constructeurs automobile dans leur objectif de créer un plus grand confort et plaisir de conduite. D'un point de vue émotionnel, l'habitacle est souvent perçu par le conducteur comme une "bulle de protection", ce qui renforce son sentiment de sécurité ¹. Toutefois, cette "surprotection" peut générer des artefacts. Sur les voitures haut de gamme, l'ambiance sonore est tellement feutrée qu'elle peut contribuer à la réduction de la vigilance du conducteur, en le rendant inconscient de ses responsabilités sur la route. Le silence total n'est donc pas envisageable dans une voiture. À l'extrême, les niveaux sonores trop élevés ne sont pas acceptés non plus, car ils peuvent créer une fatigue auditive. À un niveau sonore modéré, le confort auditif dépend de plusieurs facteurs, notamment le spectre des bruits dont certaines harmoniques émergentes peuvent avoir un effet négatif, en créant la gêne ou à l'inverse, un effet positif, en procurant un plaisir de conduite [44].

Types de clients L'identité sonore d'une voiture est développée en fonction du type de client pour lequel cette voiture est orientée. Les attentes et les exigences peuvent varier selon le sexe, l'âge, le type de conduite ou la nationalité du client. D'une manière générale, les plus jeunes tolèrent des niveaux sonores plus importants. Les gens qui adoptent une conduite plus sportive attendent que le bruit de leur voiture possède suffisamment de puissance et de "nervosité". Les clients haut de gamme associent à leurs berlines des bruits et des vibrations assurant une ambiance noble et feutrée, ce qui exige une maîtrise totale de l'acoustique, à l'aide de matériaux absorbants. Une autre catégorie de clients sont les gens qui passent beaucoup de temps à conduire et qui sont exposés à la longue à une fatigue. La qualité et en particulier la qualité sonore dépend fortement de la nationalité et des différences culturelles des clients. Certaines études internationales mettent en évidence une notion de qualité sonore différente, influencée par le contexte social et culturel des pays. Par exemple, des clients européens et asiatiques révèlent une sensibilité différente aux basses et hautes fréquences dans leur appréciation du bruit automobile dans l'habitacle [45].

¹Ceci permet d'expliquer en grande partie le comportement agressif de certains conducteurs au volant

4.2 Indicateurs psychoacoustiques

Les indicateurs psychoacoustiques servent à décrire les différents aspects auditifs des sons. La plupart de ces indicateurs ont été établis pour des sons théoriques comme les sinusoides, bruit blanc, mais aussi pour les sons musicaux, ou encore pour certains sons de l'environnement. J'ai choisi de présenter ici six de ces indicateurs : sonie, hauteur, acuité, timbre, rugosité et fluctuation. Même s'ils sont censés décrire des aspects différents des sons, ces indicateurs ne sont pas indépendants entre eux.

4.2.1 Sonie ("loudness")

L'intensité sonore est mesurée en décibels (dB) (SPL). Le décibel est une unité sans dimension qui exprime la pression acoustique. SPL désigne le niveau de pression acoustique (SPL - "sound pressure level"). L'intensité sonore est une grandeur objective, car elle représente le niveau sonore physique des sons. Or, ce niveau est différent du niveau perçu par l'oreille. Le niveau sonore perçu est une grandeur subjective et est représenté par la sonie.

Définition La sonie est l'indicateur psychoacoustique qui mesure la sensation subjective de l'intensité sonore. La sensation subjective est différente de l'intensité objective, exprimée en dB. Cette différence provient du fait, que l'oreille humaine ne traite pas de la même façon toutes les fréquences de la plage audible. La sonie dépend donc de l'intensité, mais aussi de la fréquence. Cette observation a été faite sur des sons purs par Fletcher et Munson en 1933, amenant à la définition des courbes d'isotonie [46].

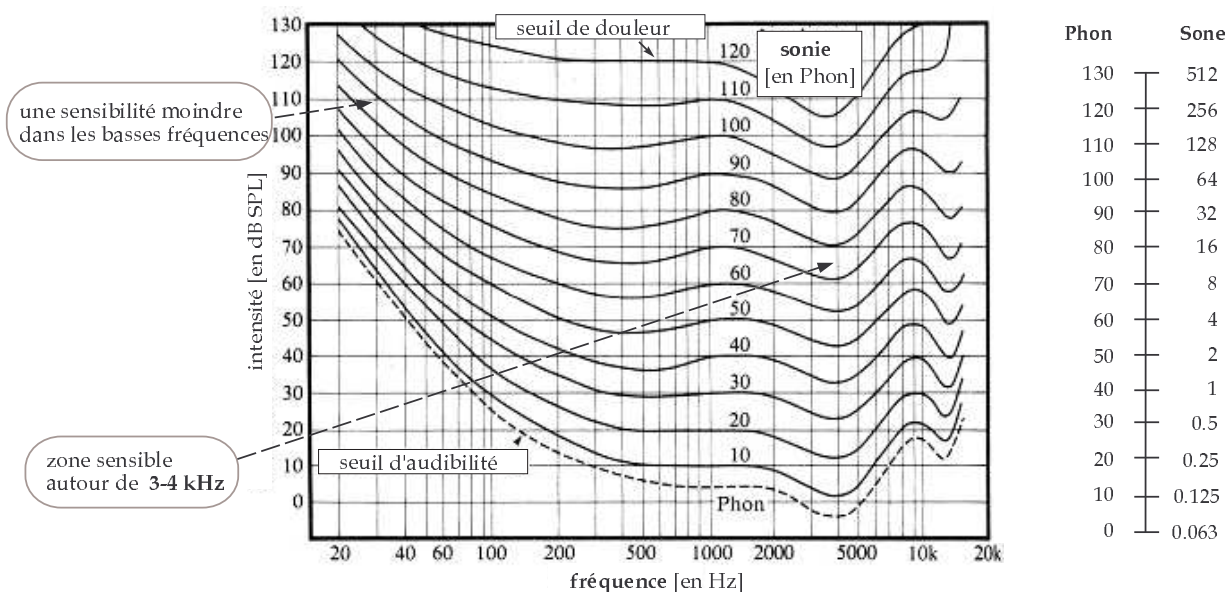


FIG. 4.4 – Les courbes d'isotonie selon Fletcher et Munson (adaptée de [47])

Les courbes d'isotonie représentent des lignes où l'intensité perçue des sons purs reste constante, pour une variation de la fréquence et l'intensité objective (**Fig. 4.4**). L'unité de la sonie ici est le Phone, alors que la fréquence et l'intensité s'expriment respectivement en Hz et dB SPL. Les courbes d'isotonie témoignent que la sensibilité de l'oreille est plus faible dans les basses que dans les hautes fréquences. En autre termes, il est nécessaire d'appliquer une intensité en dB beaucoup plus importante dans les basses fréquences afin d'obtenir la même sensation de niveau sonore que procure un son situé dans les hautes fréquences. Par ailleurs, la zone 3-4 kHz est la plus sensible pour l'oreille, car elle correspond à la résonance du conduit auditif.

Unité de sonie Les courbes d'isotonie utilisent l'unité de Phone. D'après la définition, 1 Phone correspond à l'intensité de 1 dB SPL pour la fréquence 1 kHz. En fait, le Phone est une unité de mesure, utilisée seulement dans l'objectif de comparer des sons possédant la même sonie. On ne peut pas utiliser l'unité Phone pour comparer des sons avec une sonie différente. Par exemple, on ne peut pas dire qu'un son de 40 Phones est perçu deux fois plus fort qu'un son de 20 Phones. Afin de permettre la comparaison des sons avec une sonie différente, une autre unité est introduite, le Sone. Le Sone a été introduit par

Stevens en 1936 et est défini de manière que 1 Sone soit équivalent à 40 Phones, quelque soit la fréquence. La relation entre le Phone et le Sone s'exprime comme

$$\text{Phone} = 40 + 10 \cdot \log_2(\text{Sone}). \quad (4.1)$$

Cette relation est définie ainsi pour que l'unité de Sone puisse suivre la logique de l'échelle des dB. A chaque fois qu'on double les Sones, ceci correspond à un niveau sonore perçu deux fois plus fort. Deux Sones correspondent à 50 Phones, quatre Sones à 60 Phones, etc (**Fig.4.4 (à droite)**).

Sonie et niveaux pondérés Les sons n'ont pas la même sonie en fonction de la fréquence. Pour deux sons, par exemple de 100 Hz et 1 kHz, qui ont le même niveau sonore de 50 dB, le son à 1 kHz va produire une intensité sonore perçue beaucoup plus forte que le son à 100 Hz. Pour tenir compte de cette particularité de l'oreille, on utilise couramment des filtres qui pondèrent les niveaux en fonction des fréquences à partir des courbes d'isonie. Les trois types de pondération utilisée le plus couramment sont les pondérations A, B et C, notées respectivement dB(A), dB(B) et dB(C). Ces trois filtres suivent les courbes d'isonie respectivement pour 40, 70 et 100 Phones.

Les niveaux pondérés s'avèrent utiles comme une première approche dans l'estimation du niveau sonore perçu. Toutefois, ils ne fournissent pas une mesure précise de la sonie. En fait, l'utilisation du niveau pondéré A (équivalent à 40 Phones) pour des signaux qui ont un niveau sonore plus fort, peut conduire à une sous-estimation des basses fréquences. Ceci provient des profils des courbes d'isonie qui s'aplatissent avec l'augmentation du niveau. De plus, les niveaux pondérés ne tiennent pas compte d'un mécanisme important de l'oreille, qui influence la perception de niveau sonore. Ce mécanisme est le masquage fréquentiel et il se traduit par le fait que certaines fréquences se retrouvent masquées par d'autres et par conséquent, ne contribuent pas à l'estimation de la sonie.

Bandes de Bark et sonie Le mécanisme du masquage est illustré à travers la représentation du système auditif comme un banc de filtres. Ces filtres expriment le fait que l'oreille traite différemment les fréquences. La plage audible est donc découpée en 24 bandes, appelées bandes critiques ou bandes de Bark. Ces bandes ont une largeur constante jusqu'à 500 Hz et au-delà, leur largeur augmente d'environ 20 % avec la fréquence, ce qui est relativement proche du tiers d'octave (26 %). Le calcul de la sonie globale se fait en estimant d'abord une grandeur, appelée sonie spécifique et associée à chaque bande critique. Deux contributions sont à considérer dans le calcul de la sonie spécifique. La première est la sonie dite "de cœur", qui est due directement des composantes spectrales contenues dans la bande critique. La deuxième contribution est la sonie des flancs qui tient compte des bandes voisines contribuant également dans la perception du niveau. La sonie du cœur et la sonie des flancs forment un motif d'excitation (en anglais - "excitation pattern") qui est à la base du calcul de la sonie.

Modèles de sonie Le premier modèle de sonie, proposé par Zwicker, consiste à intégrer la sonie spécifique (notée $N'(z)$) pour toutes les bandes de Bark, afin d'obtenir la sonie globale N

$$N = \int_0^{24} N'(z) \cdot dz \quad (4.2)$$

Le modèle de sonie de Zwicker est adapté aux sons stationnaires et est à la base de la norme *ISO532* de 1975 [48]. Or, pour les sons non-stationnaires, la sonie dépend aussi de la durée. Afin de s'adapter aux signaux qui varient avec le temps, Zwicker propose plus tard une correction du modèle de la sonie [49]. Le modèle de sonie est aussi revisité par Moore et Glasberg qui proposent une nouvelle façon de calculer les filtres auditifs et le pattern d'excitation [50]. Un exemple de la sonie spécifique de deux quarts de cycle moteur est présenté sur la **Fig. 4.5**. On peut observer que les explosions dans les cylindres possèdent un niveau subjectif différent les uns des autres, ce qui dépend fortement de leur répartition spectrale, exprimée à travers l'échelle des Barks. Cet exemple est détaillé d'une manière plus approfondie dans la troisième partie de la thèse.

Les modèles de sonie existants n'intègrent que partiellement des informations comme la perception fréquentielle et le masquage, la localisation dans l'espace et certaines spécificités de la perception temporelle des sons. La difficulté de mettre en œuvre un modèle de sonie pour les sons complexes vient du fait qu'il doit tenir compte de plusieurs phénomènes auditifs. Un problème particulier représente le calcul de la sonie des sons de très courte durée dits impulsifs, très répandus en milieu industriel [51].

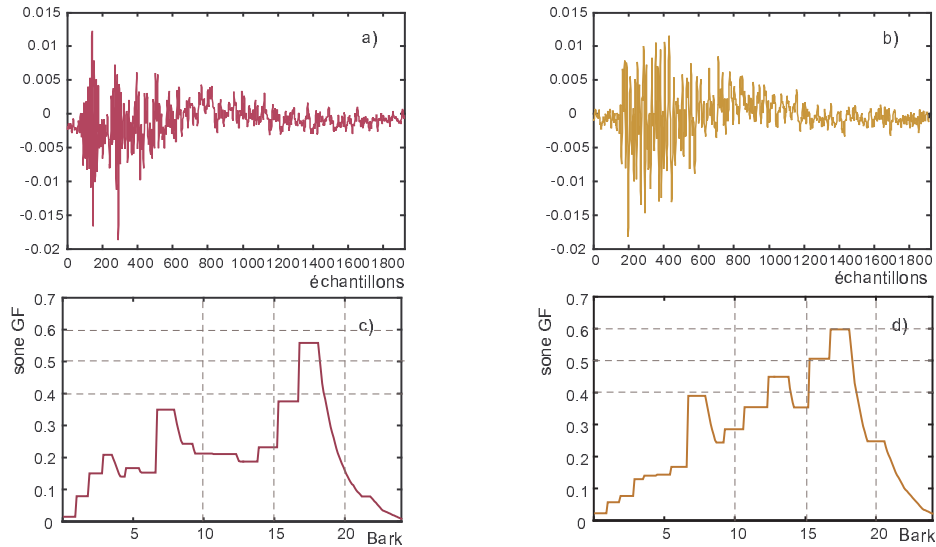


FIG. 4.5 – Deux quarts de cycle moteur au ralenti (a,b) et leurs sonies spécifiques respectives (c,d)

Sonie ("loudness") et bruyance ("noisiness") La sonie correspond à la sensation de force sonore. Toutefois, il existe un attribut, la bruyance, qui transcrit l'intensité sonore des bruits. La sonie et la bruyance semblent être deux attributs différents, comme le confirment des études d'Izumi sur les bruits impulsifs [52]. Les différences systématiques entre les deux indicateurs laisse supposer qu'il s'agit bien de deux attributs sonores distincts. Ce résultat est essentiellement valable pour les sons impulsifs. Toutefois, le paramètre "noisiness" semble ne pas avoir une grande popularité comme indicateur perceptif.

4.2.2 Hauteur ("pitch")

La hauteur est un attribut auditif qui traduit la sensation de "grave-aigu". Toutefois, derrière cette simple définition se cachent plusieurs questions perceptives.

Hauteur tonale et chroma La hauteur tonale, désignée en anglais comme "pitch", représente une facette de la hauteur qui traduit la sensation "aigu-grave" en fonction de la fréquence. Toutefois, la hauteur tonale est caractérisée aussi par une autre dimension, appelée "chroma". La chroma rapproche les sons qui ont une octave de différence. Par exemple, les sons "la" à 220 Hz et "la" de 440 Hz ont un écart d'une octave et sont perçus plus proches en hauteur, que les sons "la" à 220 Hz et le "si" à 247 Hz qui ont qu'un ton d'écart. Pour des sons complexes harmoniques, la hauteur tonale est principalement influencée par la fréquence fondamentale.

Hauteur des sons complexes Tandis que la hauteur tonale des sons purs est associée à une seule fréquence dans le spectre, la hauteur perçue des sons complexes est moins bien déterminée. En raison de leur spectre très riche, l'estimation de la hauteur des sons complexes est grandement influencée par leur timbre. Afin de faire la distinction entre ses différents aspects, la hauteur est décrite à l'aide de trois grandeurs : la hauteur fondamentale, la hauteur brute et la hauteur spectrale. La hauteur fondamentale est liée à la fréquence fondamentale du son complexe périodique et l'appellation anglaise est "low pitch" ou "complex pitch". La hauteur brute est une grandeur, introduite par Risset, qui s'oppose à la hauteur tonale et tient compte du timbre des sons [?]. Il est montré que certains sujets sont plus sensibles à la hauteur brute qu'à la hauteur tonale. Quant à la hauteur spectrale, elle est liée à la position du centre de gravité des composantes du spectre, i.e à la brillance des sons, mesurée à l'aide de l'indicateur de l'acuité.

Phénomènes liés à la hauteur Un phénomène intéressant lié à la hauteur est la fondamentale absente. En absence de la fréquence fondamentale, l'oreille peut reconstruire une sensation de hauteur fondamentale, à l'aide seulement des harmoniques de rang plus élevés dans le spectre. Ce phénomène a été observé non seulement chez un auditeur adulte mais également chez des nourissons de 8 mois, ainsi que certains animaux. La perception de la hauteur est aussi un phénomène culturel et musical. Selon Risset, les musiciens sont très sensibles à la hauteur fondamentale, alors que les non-musiciens sont attachés à la hauteur brute.

4.2.3 Acuité ("sharpness")

La notion d'acuité comme un attribut psychoacoustique est introduite par Bismark au début des années 1970, lors d'une série d'études sur le timbre des sons [53]. Cet attribut sonore est issu de la verbalisation du timbre de différents sons et exprime la sensation de brillance d'un son.

Définition L'acuité est une indication de l'équilibrage spectral entre les basses et hautes fréquences. Verbalement, elle peut se traduire par l'opposition de "mat-brillant". L'acuité d'un son augmente au fur et à mesure que les hautes fréquences deviennent plus importantes dans le contenu spectral du signal. Donc il existe une forte corrélation entre la hauteur tonale (dans ses limites de perception, jusqu'à 5-6 kHz) et l'acuité. Toutefois, le terme "aigu" n'exprime pas très précisément l'acuité, car elle a plutôt un sens de concentration spectrale. Perceptivement, les sons possédant un spectre prépondérant dans les hautes fréquences donnent une sensation de "concentration" et de "perçant", alors que les sons avec un spectre grave, donnent une impression de "volume" et d'"ampleur".

Acuité et agrément L'intérêt envers l'acuité est manifesté très fortement à la suite d'une étude de Terhardt et Stoll effectuée en 1981 sur l'agrément (en anglais - "pleasantness") [54]. Les auteurs utilisent des sons de l'environnement comme des bruits de moto, des sonneries de téléphone, des bruits d'aspirateur ou encore la voix humaine. L'étude déduit une forte corrélation entre l'acuité et l'agrément au sens où l'augmentation de l'acuité provoque la forte baisse de l'agrément. Ce résultat met en évidence l'importance de l'acuité dans l'évaluation de la qualité sonore. L'acuité s'oppose également à la notion de consonance.

Mesure de l'acuité Une approximation simple pour mesurer l'acuité peut être donnée par le barycentre du spectre. La mesure plus précise prend en compte les bandes critiques. La première formulation pour mesurer l'acuité est faite par Bismarck et est ensuite reprise et ajustée par Aures [55]. Aures introduit aussi l'unité de l'acuité, l'acum. La mesure de 1 acum représente l'acuité du signal de référence, un bruit bande étroite, centré autour de 1 kHz, avec un niveau de 60 dB de la bande critique. La formule d'acuité finalement retenue est donnée par Zwicker et Fastl [36]. Elle s'inspire du calcul de la sonie, et fournit l'acuité en fonction de la sonie spécifique N'

$$S = 0.11 \frac{\int_0^{24Bark} N' g(z) z dz}{\int_0^{24Bark} N' dz}, \quad (4.3)$$

où $g(z)$ est un facteur de pondération, associé à la bande critique. Le facteur de pondération reste égal à 1 jusqu'à la 16-ème bande de Bark et croît au-delà pour devenir égal à 4 pour la 24-ème bande.

"Sharpness" et "booming" Un attribut psychoacoustique souvent utilisé dans la description des bruits moteur est désigné en anglais comme "booming", traduit à peu près comme bourdonnement. Le booming est en quelque sorte un attribut complémentaire de l'acuité. En effet, si l'acuité est une indication de la prépondérance des composantes hautes fréquences dans le spectre du signal, le bourdonnement est une mesure des composantes basses fréquences dans le spectre. Plus particulièrement, le bourdonnement est associé à l'évolution de la fondamentale moteur dont la plage fréquentielle ne dépasse pas 250 Hz. D'une manière similaire à l'acuité, le bourdonnement peut être calculé en utilisant une fonction de pondération qui privilégie les basses fréquences.

4.2.4 Timbre

Le timbre est une des caractéristiques les plus importantes des sons, il est en quelque sorte leur "carte d'identité". Grâce au timbre on peut identifier la source sonore correspondante à un son.

Définition Classiquement, la norme ANSI de 1960 définit le timbre comme la propriété perceptive qui différencie deux sons, possédant la même sonie et hauteur [56]. Une même note, jouée avec la même force sonore par un piano et une guitare, n'aura pas les mêmes propriétés sonores, car les deux instruments ne possèdent pas la même sonorité. Toutefois, la définition classique du timbre ne prend pas en compte la durée des sons et pourtant c'est un paramètre très important. Or, pour comparer les timbres de deux sons, ils doivent être de durée égale.

Influences spectrale et temporelle Le timbre dépend à la fois des paramètres spectraux et temporels. L'enveloppe spectrale joue un rôle important, car elle détermine la richesse du spectre et la

répartition énergétique entre les partiels. La hauteur dépend globalement de la fondamentale et les premiers harmoniques, mais c'est le reste du spectre qui apporte la différence entre les sonorités. D'un autre côté, l'enveloppe temporelle est très importante, car elle comporte entre autres l'attaque des sons. Or, si on enlève les premières millisecondes qui contiennent l'attaque d'un son de piano, l'instrument n'est plus reconnaissable. Une autre expérience de Risset montre que si le son est joué à l'envers, tout en conservant son spectre, ceci apporte un changement radical dans le timbre [57]. Par ailleurs, il est estimé qu'une durée minimale de 40-50 ms est nécessaire pour que l'oreille puisse percevoir le timbre d'un son.

Dimensions du timbre Le timbre est un attribut psychoacoustique multi-dimensionnel. En comparaison, la sonie peut être attachée à une seule caractéristique sonore - la force sonore. La difficulté de mesurer le timbre a amené les chercheurs à décomposer l'espace du timbre en plusieurs dimensions. Cette démarche est effectuée à l'aide d'une mesure de similarité entre les sons, appelée analyse multi-dimensionnelle (MDS : Multi-dimensional scaling), qu'on utilisera dans la troisième partie de la thèse, consacrée aux tests perceptifs. Une étude importante de Grey sur les sons des instruments de musique permet d'identifier 3 dimensions sous-jacentes du timbre [58]. La première dimension est associée à la répartition spectrale, mesurable par l'acuité. La deuxième dimension correspond à la vitesse de décroissance des différents harmoniques du spectre des sons. Enfin, la troisième dimension semble reliée à l'attaque des sons. Cependant, il faut garder à l'esprit que le même si le timbre est censé être indépendant de la sonie et la hauteur, il peut influencer leur perception.

4.2.5 Rugosité ("roughness")

Le phénomène de rugosité a été démontré en premier par Helmholtz en 1877, l'expérience consistant à jouer simultanément deux sons purs. Les fréquences proches des sons produisent dans certains cas des battements, et créent ainsi une sensation de rugosité sonore. La rugosité peut être provoquée aussi par un son, modulé en amplitude, comme le montrent Terhardt [59] et Aures [55].

Interprétation Selon Pressnitzer, la rugosité représente un attribut psychoacoustique de base du système auditif, tout comme la sonie, et la hauteur [60]. Le terme "rugosité" au sens auditif est souvent difficile à décrire. Toutefois, la rugosité peut être considérée comme la sensation inverse de la consonance. Le son de sifflet à bille peut illustrer l'idée d'un son rugueux. Le fait que la rugosité peut être observée à la fois à partir du rapprochement des fréquences de deux sons purs et à partir d'une modulation en amplitude, détermine deux approches de la rugosité : l'approche spectrale et l'approche temporelle.

Approche spectrale L'approche spectrale de la rugosité est fondée sur l'étude de l'écart fréquentiel entre deux sons purs. La forme d'onde d'un signal constitué de deux sinusoides peut également s'écrire en termes de l'écart fréquentiel Δf et de la fréquence moyenne f_{moy} ,

$$\begin{aligned} s(t) &= \cos 2\pi f_1 t + \cos 2\pi f_2 t, & \Delta f &= f_2 - f_1, \\ \Updownarrow & & \text{avec} & \\ s(t) &= 2 \cos 2\pi \frac{\Delta f}{2} t \cdot \cos 2\pi f_2 t, & f_{moy} &= \frac{f_1 + f_2}{2}, \end{aligned} \quad (4.4)$$

Cette écriture permet l'interprétation du signal comme la modulation d'une sinusoïde de fréquence f_{moy} , appelée fréquence porteuse, avec une fréquence $\frac{\Delta f}{2}$, appelée fréquence modulante. L'approche spectrale permet de mettre en évidence la relation entre la rugosité et les bandes critiques. Plus précisément, la rugosité provoquée par deux sons purs dépend du rapport

$$\frac{\Delta f}{\text{largeur de bande critique à } f_{moy}} \quad (4.5)$$

Pressnitzer montre que la rugosité disparaît nettement lorsque l'écart fréquentiel est juste supérieur à une bande critique, par exemple si l'écart représente 110 % de la largeur de la bande [60]. Si l'écart fréquentiel est encore plus important, on entend deux sons séparés. Cette conclusion est en accord avec le seuil de consonance, estimé par Plomp et Levelt [61].

Approche temporelle L'approche temporelle de la rugosité permet de décrire la transition entre trois phénomènes perceptifs : la fluctuation, la rugosité et la perception de composantes séparées. Un son modulé en amplitude (**Fig.12.15**), comportant une fréquence porteuse f_c , une fréquence modulante f_m et une profondeur de modulation m , s'écrit

$$s(t) = (1 + m \cdot \cos(2\pi f_m t)) \cdot \cos 2\pi f_c t \quad (4.6)$$

A une fréquence de modulation basse, on perçoit nettement des fluctuations. Au moment où la fréquence de modulation devient trop élevée, l'oreille ne peut plus suivre la succession de maxima et minima du niveau de l'amplitude, et la sensation de rugosité apparaît. En augmentant davantage la fréquence de modulation, la rugosité disparaît pour laisser place à un son modulé et deux composantes tonales.

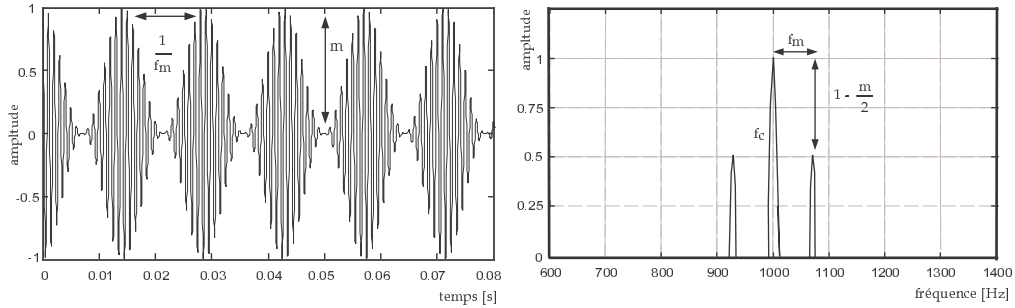


FIG. 4.6 – L'approche temporelle de la rugosité : un son modulé en amplitude

Paramètres influants sur la rugosité L'approche temporelle fournit plusieurs pistes pour étudier la rugosité. Elle donne, entre autres, des paramètres principaux qui influencent la rugosité des sons purs : la fréquence modulante, le degré de modulation, le rapport entre la fréquence modulante et la fréquence porteuse, la phase, l'enveloppe temporelle et le niveau.

La fréquence de modulation au-dessous de 20 Hz provoque plutôt une sensation de fluctuation que de rugosité. A partir de 20-30 Hz de fréquence modulante, la rugosité augmente avec la fréquence, pour atteindre un maximum, qui dépend également de la fréquence porteuse. Au moment où la fréquence de modulation dépasse une certaine limite, comprise entre 100 Hz et 250 Hz, la rugosité disparaît et on perçoit des composantes spectrales distinctes.

Le degré (ou encore profondeur) de modulation varie entre 0 et 1, et pour $m = 1$ on dit que la modulation est totale. La rugosité atteint un maximum pour une valeur de m situé au milieu de l'intervalle. Selon Terhardt, la rugosité est proportionnelle (en première approximation) au second degré de la profondeur de modulation (i.e $r \sim m^2$) [59]. D'autres auteurs trouvent un exposant de 1,4 et pas 2.

L'influence de l'enveloppe temporelle sur la rugosité est étudiée par Pressnitzer [60]. Il utilise deux signaux dont la forme d'enveloppe est respectivement en "dent de scie" et en "dent de scie inversée" et démontre que ces sons, dont la valeur RMS de l'enveloppe est la même, diffèrent en rugosité. Or, les modèles de rugosité ne prennent pas en compte l'enveloppe temporelle. Les modèles spectraux utilisent seulement le spectre du signal, alors que les modèles temporels considèrent seulement la fluctuation de l'enveloppe temporelle, au sens de la valeur RMS. L'observation de l'influence de l'enveloppe temporelle sur la rugosité met en évidence l'incomplétude des modèles de rugosité existants.

Unité et modèles de rugosité L'unité de mesure de la rugosité, l'Asper, est définie par Terhardt en 1968 suite à ses études sur les sons modulés en amplitude [59]. La rugosité vaut 1 Asper pour un son de 1 kHz à 60 dB, avec une fréquence modulante $f_m = 70 \text{ Hz}$ et un degré de modulation $m = 1$.

Les deux modèles les plus connus pour le calcul de la rugosité sont le modèle d'Aures [55] et le modèle de Daniel et Weber [62]. Ces modèles, basés sur l'approche temporelle de la rugosité, calculent la rugosité globale en additionnant les rugosités partielles dans chaque bande critique. Sur ce dernier point, les modèles de rugosité s'inspirent du modèle de Zwicker de calcul de la sonie. Toutefois, il est difficile de construire un modèle élaboré de calcul de rugosité, au vu des nombreux paramètres à prendre en compte et de la complexité des sons réels.

Les modèles de rugosité supposent globalement deux propriétés : l'additivité des rugosités partielles et le fait que les régions fréquentiellement éloignées ne contribuent pas à la rugosité globale. Ces deux suppositions sont mises en cause dans le cas des bruits large bande. Par ailleurs, un modèle de rugosité doit également intégrer, comme le suggère Pressnitzer, l'information de l'enveloppe temporelle du signal.

Rugosité des bruits moteur La rugosité est un attribut psychoacoustique qui fait l'objet d'un grand intérêt dans le contexte automobile, du fait que l'étude de la rugosité est souvent associée au caractère "sportif" des bruits moteur. Le spectre très large bande des bruits moteur rend difficile l'estimation de la rugosité totale. Il n'est pas non plus possible de prédire la manière dont le système auditif intègre les rugosités partielles. La présence d'un grand nombre d'harmoniques et de demi-harmoniques détermine

une interaction très complexe dans la création de la sensation de rugosité. L'existence d'une superposition des phénomènes de rugosité et de fluctuation est même probable.

4.2.6 Fluctuation ("fluctuation strength")

La fluctuation (force de fluctuation ou fluctuation de sonie) a été introduite, parallèlement à la rugosité, dans les travaux de Terhardt [59]. La fluctuation est une quantité perceptive qui décrit les modulations d'amplitude lentes d'un signal. Une fluctuation d'amplitude basse fréquence, au-dessous de 20 Hz, rend perceptible tout un événement temporel, car l'oreille peut suivre les changements d'amplitude. La modulation de 20 Hz est considéré comme un seuil de transition entre la fluctuation et la rugosité [36].

Exemples de fluctuation L'exemple le plus naturel pour illustrer la fluctuation est celui de la parole. En effet, une fluctuation de 4-5 Hz est typique pour la parole, ce qui permet de lui associer un rythme caractéristique. Les études psychoacoustiques montrent que la plus grande sensibilité pour l'oreille est atteinte pour une fluctuation de 4 Hz. C'est la fréquence à laquelle sont réglés les signaux d'alarme. Par coïncidence, les bruits moteur au ralenti possèdent exactement la même fluctuation, d'environ 4 Hz. Ceci rend cet indicateur psychoacoustique important pour l'étude du bruit moteur au ralenti et peut éventuellement expliquer une perception plutôt temporelle des bruits au ralenti.

Facteurs influant la fluctuation D'une manière similaire à la rugosité, la fluctuation peut être introduite sous l'angle de deux approches, temporelle ou spectrale. L'approche temporelle est largement plus étudiée et détermine les facteurs influant la fluctuation : la fréquence de modulation, la profondeur de modulation, la fréquence porteuse et le niveau. La fluctuation maximale est atteinte pour une valeur de 4 Hz de la fréquence de modulation. Au-dessous et au-dessus de cette limite, la fluctuation perçue décroît. L'augmentation de la profondeur de modulation provoque une augmentation de la fluctuation, aussi bien pour les sons purs [59] que pour les bruits large bande [63].

Unité de mesure L'unité de la fluctuation qui est le vacil, où 1 Vacil correspond à un son pur de 1 kHz, modulé à 100 % avec une amplitude de modulation de 4 Hz et un niveau sonore de 60 dB. Cette unité est désignée par Fastl qui propose un modèle de la fluctuation, étroitement lié aux phénomènes de masquage temporel [63]. Une approximation pour la fluctuation F des sons purs modulés en fréquence ou en amplitude est donnée en termes de bandes critiques

$$F = \frac{0.008 \int_0^{24Bark} \Delta L/dB \text{ Bark } dz}{\frac{f_{mod}}{4Hz} + \frac{4Hz}{f_{mod}}}, \quad (4.7)$$

où ΔL se déduit des différentes courbes de masquage.

Les indicateurs psychoacoustiques qui viennent d'être exposés sont extraits de tests psychoacoustiques. Il existe plusieurs types de tests perceptifs, en fonction de l'information perceptive qu'on veut obtenir des auditeurs.

4.3 Méthodologie des tests perceptifs

Deux approches principales existent dans les tests perceptifs - verbale et non-verbale. L'approche verbale utilise les mots du langage pour décrire les caractéristiques auditifs des sons et aboutit à une sorte de portrait sonore. Dans cette catégorie, on peut citer le test des différentiels sémantiques et le test de verbalisation. Au contraire, l'approche non-verbale ne recourt pas au langage. Dans cette catégorie, on peut citer le test d'estimation de grandeur et le test par paires.

L'avantage d'une méthode verbale est la possibilité d'obtenir une information complexe qui peut révéler également certains facteurs émotionnels du jugement. L'inconvénient est l'ambiguïté qui peut exister lors de l'interprétation des mots. L'avantage des méthodes non-verbales est la plus grande précision qui ne fait pas intervenir le langage. Leur inconvénient est la difficulté d'interpréter les critères que les auditeurs ont choisi dans leur évaluation. Pour ces raisons et afin de garantir une entité des résultats, dans la pratique, une combinaison entre les méthodes verbales et non-verbales est souvent utilisée.

4.3.1 Test d'estimation de grandeur

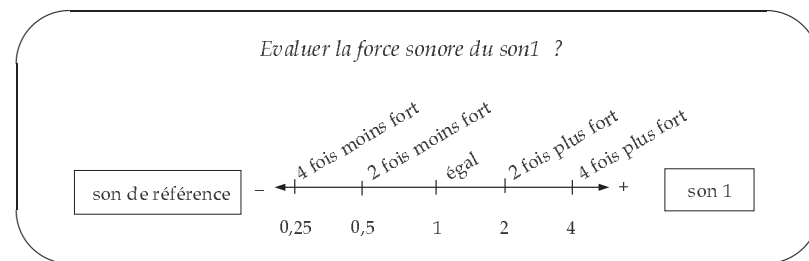


FIG. 4.7 – Schéma d'un test d'estimation de grandeur pour la sonie

Démarche Il est demandé à l'auditeur d'attribuer à chaque son une valeur proportionnelle à la sensation provoquée par un attribut sonore. Cet attribut peut être la force sonore, la hauteur ou bien l'agrément. L'estimation de grandeur peut se faire avec ou sans son de référence. Dans le cas de présence d'un son de référence, l'auditeur doit estimer sa sensation en termes de 2 fois ou 4 fois plus ou moins fort par rapport à la référence (**Fig.4.7**). On obtient ainsi une échelle de rapport de sensations.

Remarques Ce type de test a été largement utilisé dans les études psychoacoustiques sur des sons et des bruits théoriques. Il est moins souvent appliqué sur des sons réels.

4.3.2 Test d'évaluation

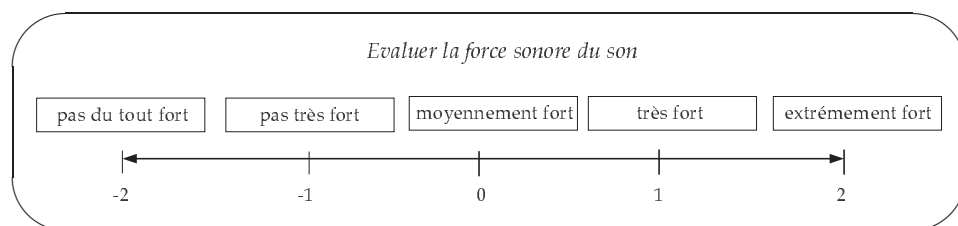


FIG. 4.8 – Schéma d'un test d'évaluation de la sonie

Démarche L'auditeur doit évaluer un caractère sonore, par exemple : "Est-ce que le son est agréable/fort/..?" La réponse est donnée sur une échelle reliant les notions extrêmes de l'attribut (**Fig. 4.8**).

Remarques Le test demande la participation d'au moins une trentaine de personnes. L'ordre de présentation des sons doit être aléatoire et une éventuelle répétition des sons est souhaitable afin de vérifier la cohérence des réponses.

4.3.3 Test d'évaluation comparée

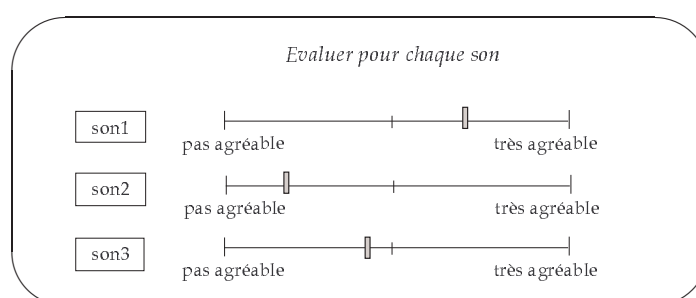


FIG. 4.9 – Test d'évaluation comparée sur l'agrément de trois sons

Démarche La démarche est semblable au test d'évaluation, sauf que dans le test d'évaluation comparée, il s'agit d'évaluer seulement une caractéristique sonore pour plusieurs sons à la fois (**Fig. 4.9**). En

effet, il s'agit à la fois d'évaluation et de comparaison.

Remarques L'avantage de cette méthode réside dans le fait qu'elle permet une comparaison plus aisée que le test par paires, tout en fournissant une très bonne précision. Toutefois, comme la plupart des méthodes utilisant une échelle d'évaluation continue, le risque existe que l'auditeur n'utilise qu'une partie de l'échelle ou à l'inverse, qu'il utilise exclusivement les extrémités, en donnant ainsi une réponse peu nuancée.

4.3.4 Test de différentiel sémantique

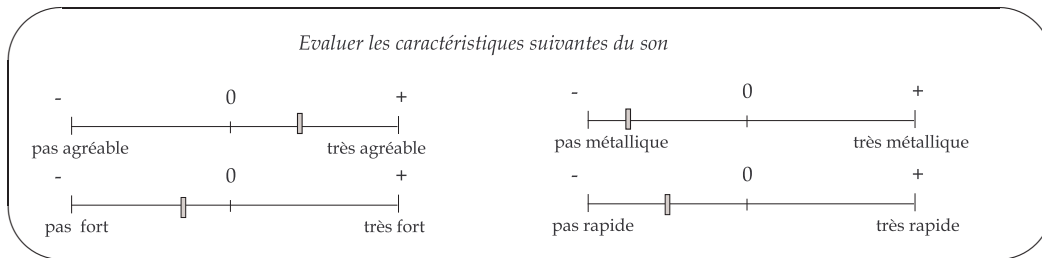


FIG. 4.10 – Test de différentiel sémantique sur quatre attributs : agréable, fort, métallique et rapide

Démarche Pour ce type de test, l'auditeur doit évaluer le son sur plusieurs échelles. Chaque échelle exprime une caractéristique spécifique du son, et varie entre les valeurs extrêmes positive et négative de cet attribut sonore (Fig. 4.10). Les caractéristiques peuvent être objectives (fort, rapide, aigu, ..) ou peuvent exprimer une échelle hédonique (agréable, désagréable ..).

Remarques Les adjectifs associés aux caractéristiques du son doivent être soigneusement choisis, afin de ne pas provoquer d'ambiguïté chez l'auditeur. Il est préférable d'utiliser un adjectif dans sa forme positive et négative et plutôt que deux adjectifs avec des sens opposés. Par exemple, un son "rugueux" est difficile à expliquer, mais reste pertinent dans l'évaluation "pas rugueux - très rugueux". En revanche, on a vu que l'opposé de rugueux au sens sonore n'est pas "lisse", comme on peut le supposer, mais comme plutôt "consonant". Ceci exprime une particularité des méthodes dites verbales, dont la méthode des différentiels sémantiques fait partie, qui est l'ambiguïté de la langue. Les auditeurs peuvent interpréter les descripteurs sonores d'une manière non unique et les adjectifs peuvent différer d'une langue à l'autre.

4.3.5 Test de comparaison par paires

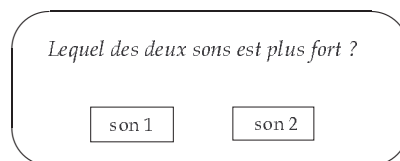


FIG. 4.11 – Test de comparaison par paire

Démarche Pour le test de comparaison par paire, l'auditeur écoute à chaque fois deux sons et doit les comparer en fonction de l'attribut sonore recherché (Fig. 4.11). La question peut demander lequel des sons est plus ou moins fort/agréable/aigu/grave etc. Ainsi sont présentées toutes les combinaisons $\frac{N*(N-1)}{2}$ par paires de l'ensemble de N sons. Les paires de sons identiques doivent aussi être présentes.

Remarques L'avantage de la méthode est la simplicité de la tâche pour l'auditeur, car il ne doit comparer que deux sons à la fois. En fait, la mémorisation de plusieurs sons est une tâche assez difficile pour la mémoire auditive. Le test de comparaison peut donner une meilleure précision, car il fournit la corrélation entre toutes les paires de sons. L'inconvénient principal de ce test est sa durée totale, due au grand nombre de combinaisons à écouter. Cependant, dans la construction des tests jouent d'autres facteurs comme la durée des sons, la difficulté de la tâche, la nature des sons. Afin d'assurer la fiabilité du test par paires, deux conditions doivent être satisfaites : que l'ordre de présentation des paires soit aléatoire et que chaque son soit présent autant de fois comme premier que comme deuxième dans la paire.

Le test de comparaison par paires permet de construire à l'aide de l'analyse en composantes principales (ACP) un espace perceptif indiquant les principaux axes pris en compte par l'auditeur dans l'évaluation.

4.3.6 Test de catégorisation (tri)

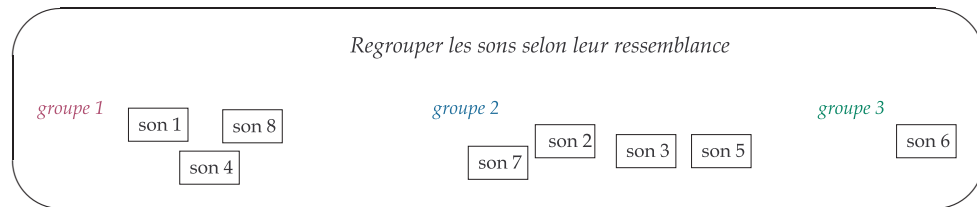


FIG. 4.12 – Test de catégorisation

Démarche La tâche de catégorisation consiste à présenter aux auditeurs un ensemble d'échantillons sonores et de demander de regrouper les sons en fonction de leur ressemblance. De cette façon, les sons qui vont ensemble sont mis dans le même groupe, et au contraire, s'ils ne sont pas jugés pareillement, sont séparés dans des groupes différents (**Fig. 4.12**). Dans le cas d'une catégorisation libre, l'auditeur a le choix de faire autant de groupes qu'il juge nécessaire. De plus, il n'est pas limité dans le nombre de sons constituant chaque groupe. Dans certains cas, le nombre de groupes peut être fixé préalablement.

Remarques La classification donne une idée sur les critères perceptifs, utilisés par les auditeurs dans l'organisation de catégories de sources et d'événement sonores dans la mémoire. Ces critères doivent être ensuite reliés avec les caractéristiques physiques des sons, afin de trouver des descripteurs appropriés pour expliquer les caractéristiques communes des sons d'un même groupe. Les résultats du test de catégorisation sont présentés dans une matrice de pseudo-distances (distances entre les groupes et pas entre les sons), où la valeur est "0" si les sons ont été regroupés ensemble et "1", le cas échéant. Le tri libre n'offre pas la précision d'un test de similarités, où les sons sont évalués plus finement par paires. Le tri libre bénéficie d'un temps d'exécution plus rapide et permet aux auditeurs de traiter un ensemble de sons plus vaste.

4.3.7 Test de description verbale

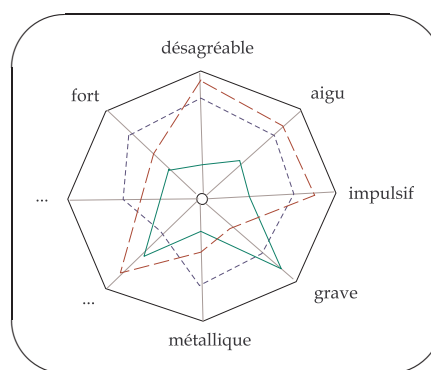


FIG. 4.13 – Espace perceptif créé à la base du test de description verbale

Démarche L'analyse de verbalisation peut être considérée comme une méthode en soi, mais elle peut être également utilisée en complément des autres test perceptifs, afin de peaufiner la signification des attributs sonores. Habituellement, cette méthode présente à l'auditeur deux sons et lui demande d'évaluer leur similarité, tout en fournissant une description complète des caractéristiques sonores qu'il perçoit (**Fig. 4.13**).

Remarques Cette méthode permet de relier l'explication détaillée des propriétés sonores avec les indicateurs psychoacoustiques. La démarche de verbalisation peut être utilisée comme une première étape de sélection des adjectifs les plus importants qui peuvent par la suite alimenter le test d'évaluation. L'inconvénient de la verbalisation est le processus relativement complexe de traitement. De plus, la richesse du langage fait que chaque auditeur va se servir de ses propres moyens d'expression. Par la suite, il est parfois difficile de faire l'équivalence des termes synonymes.

Conclusion sur la qualité sonore du bruit moteur

Ce chapitre introduit l'approche de la perception et la qualité sonore dans le contexte automobile. La notion de qualité sonore dépasse largement les limites du domaine auditif, en prenant en compte un ensemble de facteurs culturels, sociaux, émotionnels. La qualité sonore du bruit automobile est décrite sous l'angle des démarches perceptives, entreprises par les constructeurs automobiles. Ces démarches comprennent à la fois des études perceptives dans des conditions de laboratoire et des tests clients "in situ", i.e dans des conditions réelles de conduite. La démarche perceptive comporte deux facettes : la description des bruits via les indicateurs psychoacoustiques de base comme la sonie, l'acuité, la rugosité, etc, et les tests psychoacoustiques déterminant des dimensions perceptives spécifiques pour chaque ensemble de sons étudiés. L'objectif de l'étude perceptive consiste à relier ces deux facettes afin d'expliquer les dimensions perceptives en fonction des indicateurs psychoacoustiques de base. Un des objectifs de la thèse consiste à trouver des facteurs influant la qualité sonore des bruits moteur au ralenti.

Orientation de la thèse : L'étude perceptive de laboratoire est la démarche adoptée dans la thèse. Ce choix est logique, en raison de l'utilisation de signaux de synthèse, obtenus à l'aide de la méthode d'analyse-synthèse développée. Les conditions d'écoute au laboratoire donnent plus de liberté sur le choix des sons et offrent aux auditeurs la facilité de se concentrer sur les facteurs purement auditifs. Par ailleurs, le régime moteur étudié est le ralenti diesel et c'est une situation de conduite statique qui n'exige pas pour autant de prendre en compte les facteurs dynamiques de conduite. Le choix des tests perceptifs pour les bruits au ralenti est largement discuté dans la troisième partie de la thèse.

Moteur diesel au ralenti

Le ralenti moteur représente le plus bas régime auquel le moteur tourne. En fonction des moteurs, il varie entre 600 et 1000 tr/min, les valeurs les plus courantes se situant entre 750 et 800 tr/min. Après le bruit de démarrage, le ralenti est la première manifestation sonore du moteur perçue par le conducteur. N'étant pas associé à une situation réelle de conduite, ce bruit peut facilement attirer l'attention du conducteur et témoigner de la qualité sonore et du bon fonctionnement du moteur. Le bruit au ralenti peut être à l'origine d'une gêne acoustique, car le bruit de combustion y est particulièrement perceptible. En comparaison, la combustion à un régime élevé (> 3000 tr/min) peut être masquée par d'autres bruits en provenance du moteur même, du roulement ou d'air écoulé. Ce chapitre discute les facteurs influant le bruit au ralenti et introduit une approche originale de la perception rythmique du ralenti.

5.1 La combustion

La combustion est le processus majeur du cycle moteur. Elle dépend principalement de la quantité de combustible injecté, du type de l'injection (directe, indirecte, injection pilote), de la géométrie de la chambre de combustion et des réglages du fonctionnement moteur. L'étude de la combustion est importante, car elle influence directement le rendement, la durée de vie et les performances du moteur, ainsi que l'émission des gaz polluants et les caractéristiques acoustiques du moteur. L'observation de la combustion permet de détecter d'éventuelles défaillances dans le moteur.

5.1.1 Pression dans le cylindre

L'étude de la combustion passe par l'observation d'un diagramme caractéristique qui trace l'évolution de la pression des gaz en fonction de l'angle vilebrequin.

La pression dans le cylindre donne une information importante sur le processus thermique dans le cylindre et la loi d'énergie qui correspond à un type de combustion avec un réglage particulier. La courbe de pression distingue le début de l'injection et de la combustion, ainsi que le phénomène de combustion (rapide, principale, post-combustion) (**Fig.5.1**, adaptée de [7]). A la base de ces connaissances, il est possible d'effectuer une optimisation de la combustion, en vue de plusieurs objectifs : un meilleur rendu, une faible consommation de gazole, une faible émission de gaz polluants, et enfin une meilleure acoustique du moteur. Toutefois, ces améliorations ne peuvent pas être toutes atteintes à la fois.

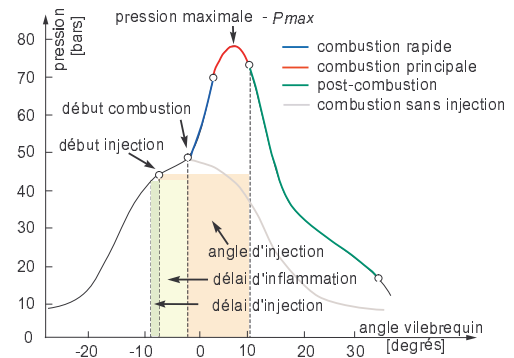


Fig. 5.1 – Pression dans le cylindre

La combustion représente l'excitation majeure de la structure du moteur. L'explosion génère une pression qui atteint 80 bars pour un moteur à essence et presque le double pour un moteur diesel. La combustion est d'une durée très courte, de l'ordre de quelques dixième de μs et s'étale spectralement sur plusieurs kHz. Pour un moteur à essence, l'étendue spectrale de l'explosion dépasse rarement 1 kHz, alors que pour un diesel, elle contient des fréquences largement plus élevées. La combustion diesel est donc un phénomène très important qui provoque un bruit impulsif large bande. La combustion est associée au bruit caractéristique de "claquement" diesel.

5.1.2 Voies de propagation

Les voies de propagation déterminent la façon dont les vibrations engendrées par la combustion vont se propager dans la structure moteur, vont la faire rayonner et générer du bruit. Les modes propres d'une structure sont associés à des fréquences propres, qui lui sont caractéristiques. Dans la propagation du bruit de combustion on distingue deux chemins de transfert : solidien, qui passe à travers la structure, et aérien, qui est dû aux résonances de la cavité d'air.

Excitation de la structure Les vibrations dues aux forces de pression, empruntent principalement trois voies de propagation dans la structure. La première passe par la tête du cylindre et la culasse et représente la transmission la plus directe. La seconde voie passe par la chemise du cylindre et la jupe du piston et représente la voie majeure de propagation des chocs du basculement du piston. La troisième voie est plutôt indirecte, car elle passe par le piston, la bielle, le vilebrequin, les paliers et enfin, le bloc cylindres. Toutefois, cette troisième voie de propagation est assez importante au-delà de 1 kHz. Localement, la culasse subit directement les forces de pression, mais l'excitation la plus importante est transmise au vilebrequin par le piston et la bielle.

Vibrations d'air Parallèlement aux vibrations de la structure, la combustion est associée aux résonances dans la chambre (et éventuellement dans la pré-chambre) de combustion. En effet, la pression exercée par le piston fait vibrer les gaz dans le cylindre. Lorsque le piston atteint le point mort haut, ces modes peuvent absorber une partie de l'énergie libérée lors de la combustion afin de faire vibrer la masse gazeuse. Dans ce processus, la pré-chambre de combustion se comporte comme un résonateur de Helmholtz et vibre à quelques kHz, alors que la chambre de combustion possède des modes de résonance plus hauts et multiples [7].

Exemple d'excitation Il est évident que la combustion est responsable d'une grande partie de l'excitation du moteur et par conséquent du bruit engendré. La différence entre un bruit avec et sans combustion d'un bruit moteur au ralenti est présentée sur (**Fig. 5.2**). Dans ce cas précis, la combustion apporte un élargissement du spectre, ainsi qu'une amplification dans la zone autour de 1 kHz. Dans le cas général, la contribution spectrale de la combustion dépend des réglages et l'architecture du moteur.

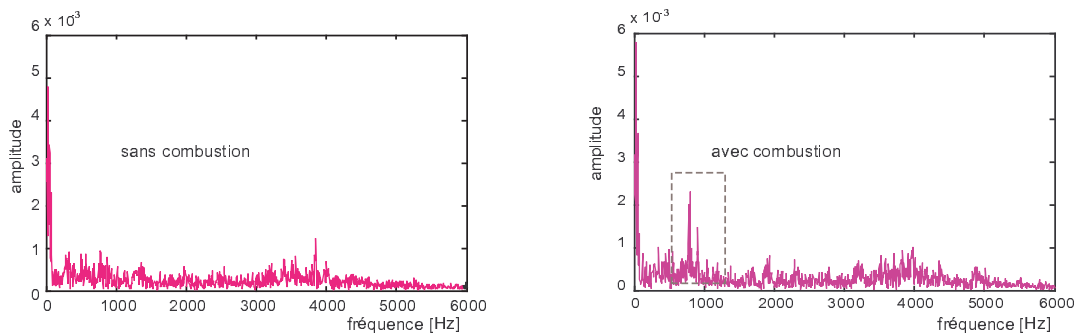


FIG. 5.2 – Spectre de bruit moteur au ralenti avec et sans combustion

Le bruit moteur avec combustion est nettement plus impulsif et "dur" à l'oreille que le bruit sans combustion. Les bruits au ralenti avec et sans combustion se trouvent respectivement sur **Piste 37** et **Piste 38**. La gêne acoustique provoquée par la combustion est également due au fait que la combustion intervient dans une zone très sensible pour l'oreille humaine (1-4 kHz), comme le témoignent les courbes d'isonomie, présentées dans le chapitre précédent.

5.1.3 Importance de l'injection

L'injection a été déjà décrite au début de cette partie, dans le chapitre consacré à l'acoustique moteur. La maîtrise de la combustion dépend grandement des techniques d'injection. Plus particulièrement, l'introduction de la technique d'injection du "common rail" permet l'amélioration de la combustion des moteurs diesel. Les systèmes les plus récents de "common rail" sont capables d'effectuer jusqu'à cinq injections par cycle, en comptant des injections pilotes, des injections principales et parfois des post-injections.

Injection pilote Comme on l'a déjà mentionné, l'injection pilote consiste à injecter une très petite quantité de gazole juste avant l'injection principale. Le résultat est une combustion moins violente, une meilleure gestion du carburant et d'un point de vue acoustique, un bruit de combustion moins impulsif. L'injection pilote permet donc de réduire la consommation de gazole, les émissions des gaz polluants, et de plus, d'améliorer la qualité du bruit émis. Le bruit des moteurs diesel possédant une injection pilote est jugé d'une qualité sonore nettement meilleure que le bruit sans injection pilote.

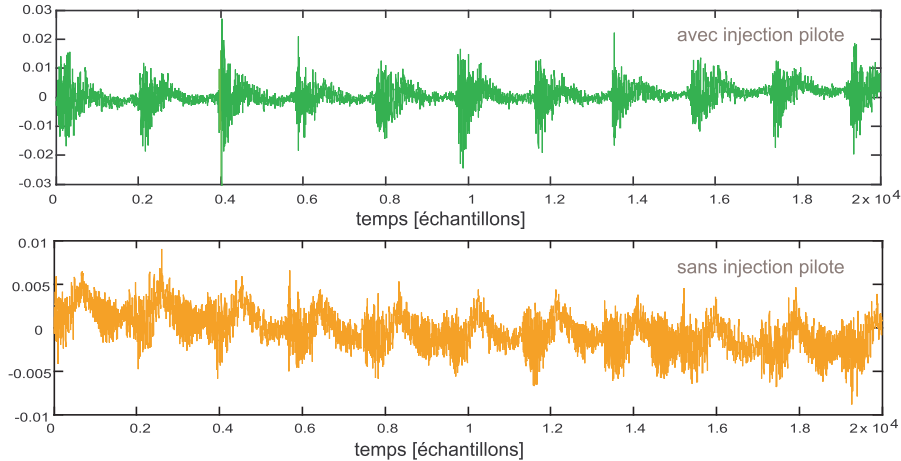


FIG. 5.3 – Signaux temporels de bruit moteur au ralenti avec et sans injection pilote

Les différences entre les enveloppes temporelles et spectrales des bruits moteur avec et sans injection pilote sont présentées respectivement sur (Fig. 5.3) et (Fig. 5.4). Il apparaît clairement que le bruit sans injection pilote possède un spectre beaucoup plus riche, surtout dans la région fréquentielle 1-4 kHz. Au contraire, le bruit avec injection pilote se caractérise avec une excitation moindre de la structure du moteur. Les signaux temporels montrent que l'absence de l'injection pilote introduit une variation très lente du signal autour de la moyenne, alors que le bruit avec injection pilote semble contribuer à un certain équilibre.

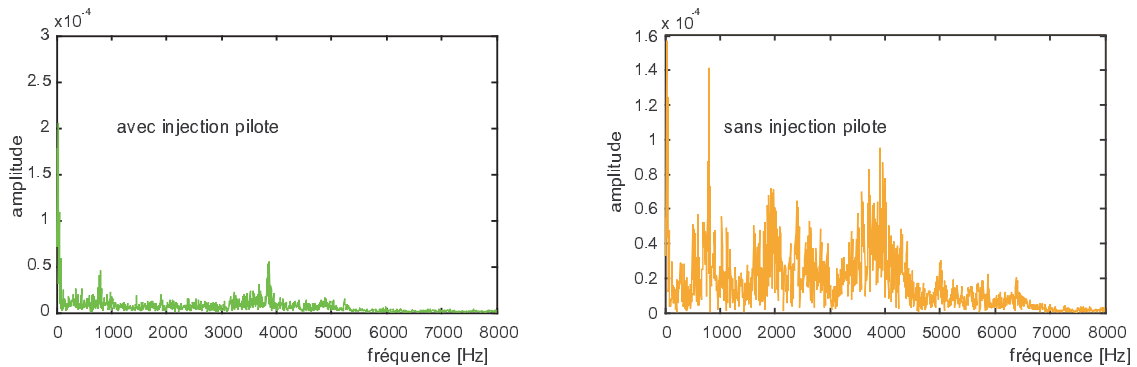


FIG. 5.4 – Spectres de bruit moteur au ralenti avec et sans injection pilote

Effectivement, à l'écoute, le bruit avec une injection pilote (**Piste 39**) possède un niveau nettement plus faible et une structure moins impulsive que le bruit sans injection pilote (**Piste 40**).

5.1.4 Qualité du bruit de combustion

La qualité du bruit de combustion fait l'objet de nombreuses études. Trois phénomènes physiques intervenant dans le cycle moteur sont mentionnés comme susceptibles de provoquer la gêne auditive. Ce sont la compression, la combustion et la résonance. La résonance correspond à une oscillation rapide des gaz autour de 6-10 kHz, due à l'augmentation de la pression dans la chambre de combustion. C'est un phénomène haute fréquence. Deux indicateurs sont déterminés à la base de la pression dans le cylindre [64]. Le premier est l'énergie de résonance, significative dans l'excitation du bloc moteur. Le deuxième est lié à la dérivée maximale de pression de la combustion, qui caractérise la brusque montée de pression au moment de la combustion principale et qui est à l'origine du "diesel knock" (claquement diesel).

Le claquement diesel est un phénomène normal pour les moteurs diesel, car il est dû à la nature de la combustion par autoinflammation. En revanche, pour les moteurs à essence, un phénomène particulier, le cliquetis (en anglais désigné seulement par "knock") est un phénomène anormal dû à un mauvais contrôle de la combustion. Provoqué par une combustion trop violente, le cliquetis est caractérisé par un bruit métallique et nuit non seulement à la qualité du bruit, mais peut aller jusqu'à l'endommagement des

parties moteur, notamment les pistons. C'est un phénomène indésirable pour le bon fonctionnement du moteur.

Le claquement sur un moteur diesel est également un phénomène qui nuit à la qualité sonore. L'excitation primaire des claquements diesel vient d'un côté de la pression dans le cylindre, provoquée par le mouvement du piston, et d'autre part, de la combustion elle-même. Ce type de bruit trouve ses origines dans l'attelage mobile et est particulièrement perceptible au bas régime (ralenti), lors d'un démarrage à froid. Le claquement diesel à froid est un bruit typique des vieux moteurs diesel.

Ralenti diesel à froid et à chaud Les moteurs diesel se distinguent des moteurs à essence surtout par leur claquement au ralenti et au démarrage à froid. Ces claquements sont dûs à la montée rapide de pression dans la chambre de combustion, ce qui provoque le basculement du piston. Les chocs du piston contre la chemise du cylindre sont propagés par la structure du moteur. Outre la combustion-même, le basculement du piston est considéré comme une des causes principales du bruit diesel.

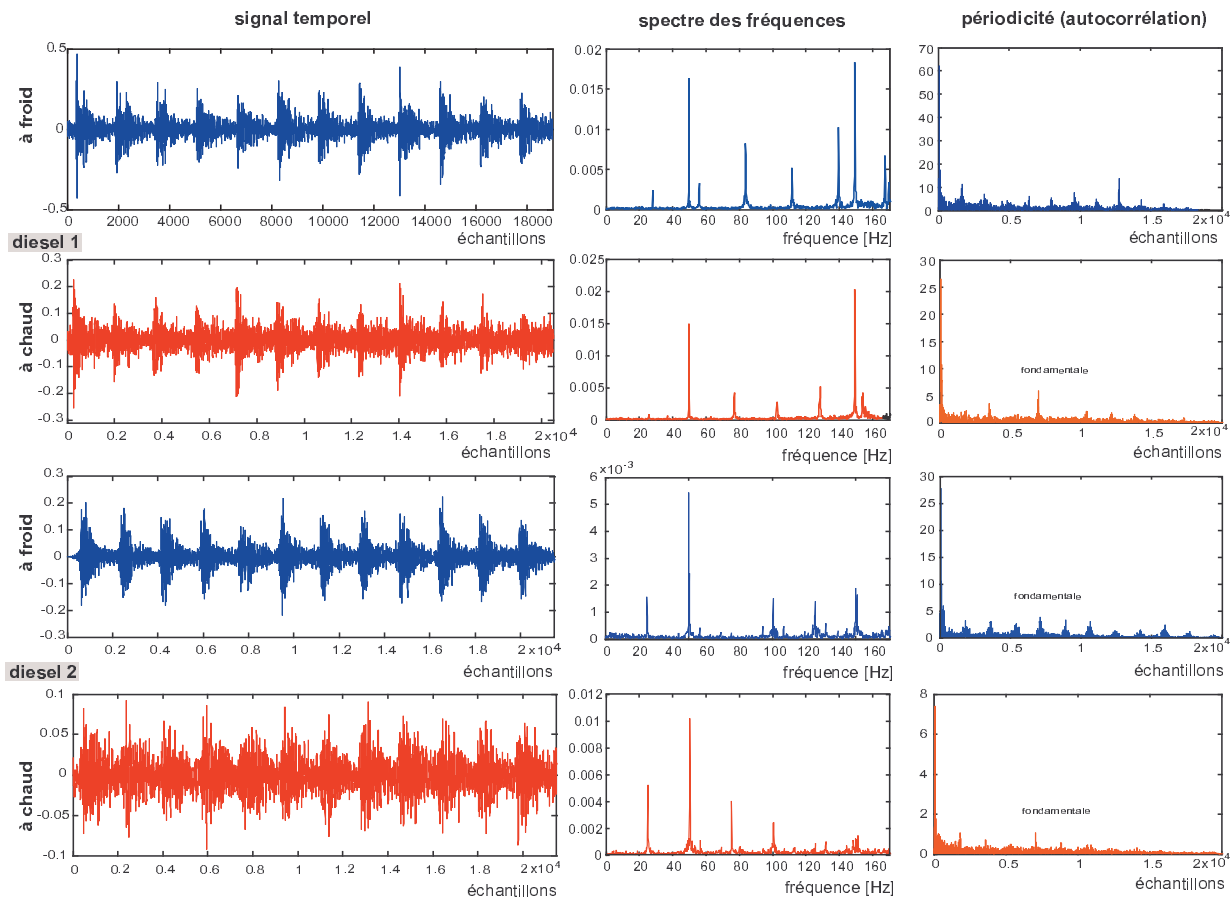


FIG. 5.5 – Deux moteurs diesel au ralenti et leur bruits respectifs à chaud et à froid

Le bruit au ralenti diesel à froid est assez différent du bruit au ralenti à chaud (**Fig.5.5**). A froid, les explosions dans les cylindres sont plus impulsives et "régulières". Cette régularité est considérée au sens où souvent un des cylindres est plus fort que les autres. Par conséquent, le bruit au ralenti à froid pour un moteur 4 cylindres se rapproche du bruit d'un monocylindre (**Piste 41**). En termes d'harmoniques moteur, ceci veut dire, que l'harmonique F0.5 est très présente. Ceci engendre une périodicité quadruple de la fondamentale moteur. Le ralenti à chaud se manifeste par un bruit moins régulier au sens où les différences entre les explosions deviennent plus importantes. Probablement, au ralenti à chaud commencent à se manifester de petits retards ou avancements des explosions qui sont dûs à la nature non déterministe des gaz dans le cylindre. Le résultat perceptif est une régularité moins exprimée qu'au ralenti à chaud. Il faut mentionner, que cette moindre régularité ne doit pas dépasser un certain seuil, auquel cas le bruit au ralenti risque d'être considéré trop irrégulier et par conséquent traduirait un dysfonctionnement du moteur.

5.2 Etude perceptive du ralenti

Une grande majorité des études perceptives du bruit moteur sont orientées vers le bruit à un régime élevé. Relativement peu d'études sont consacrées au bruit du ralenti diesel. Pourtant, ce bruit influence grandement la qualité sonore automobile. Le bruit au ralenti appartient à une classe de sons, désignés comme des bruits impulsifs répétitifs. Ces bruits sont provoqués par le fonctionnement répétitif d'un mécanisme, souvent rotatif, par exemple le bruit d'hélices d'hélicoptère, de tondeuse à gazon, de mitrailleuse.

5.2.1 Sons impulsifs

Le son impulsif est le résultat auditif d'un impact, de la mise en vibration d'un système résonnant. Le son d'une cloche, les claquement de deux mains, le bris d'un verre, un son de percussion sont des exemples de sons impulsifs, alors que les bruits de marteau-piqueur, de mitrailleuse, de moteur font partie des bruits impulsifs répétitifs. Il n'existe pas de définition rigoureuse d'un son impulsif. Cependant, comme son nom le suggère, le critère le plus important est sa durée, i.e il s'agit d'un son bref. La norme ISO fixe cette durée à une seconde [65]. Ainsi, si le son a une durée supérieure à 1 s, il est continu, si cette durée est inférieure à 1 s, il est impulsif. Les sons impulsifs plus courts que 300 ms sont dits impulsionnels [66].

Définition La durée d'un son impulsif ne suffit pas pour décrire et distinguer la variété de sons impulsifs. Une définition plus précise, donnée dans [52], considère 5 paramètres physiques du signal. Parmi ces paramètres, quatre sont issus de l'enveloppe temporelle du signal, le cinquième est lié au spectre (**Fig. 5.6**). Excepté la durée, dont l'influence a déjà été discutée, et la période dans le cas de sons impulsifs répétitifs, deux paramètres temporels majeurs sont ajoutés. Ils sont étroitement liés à l'aspect impulsif du signal. Le premier est indiqué par le facteur de crête ("crest factor") qui représente le rapport entre le maximum d'amplitude du signal lors de l'attaque et la valeur RMS du signal. Ce rapport doit être d'au moins 10 dB pour considérer le son en tant qu'impulsif. Le deuxième est le temps d'attaque qui porte une information supplémentaire sur l'impulsivité du son.

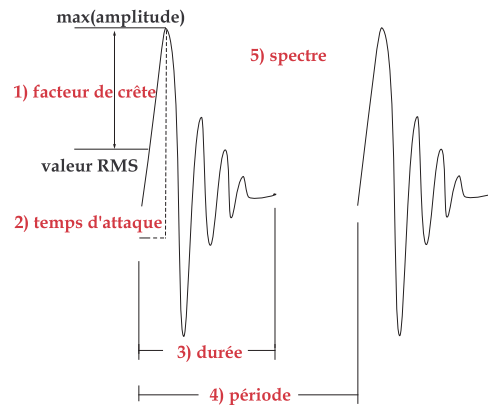


Fig. 5.6 – Représentation d'un son impulsif en fonction des paramètres du signal

Les sons impulsifs répétitifs Les sons impulsifs répétitifs représentent une grande famille de sons impulsifs. Ils sont souvent le résultat de l'activité mécanique et répétitive d'une machine. Il s'agit de bruits de marteau-piqueur, d'hélicoptère, de moto, de moteur de voitures. L'intérêt pour les sons impulsifs répétitifs est d'autant plus grand, que les sons qui provoquent le plus souvent la gêne sont néfastes pour la santé de l'homme et par conséquent sont soumis aux standards et normes pour le respect de l'environnement. Une raison importante est le niveau sonore très élevé des sons impulsifs répétitifs qui est nettement plus gênant que le niveau d'un son impulsif isolé. En effet, le degré de répétition (i.e le nombre de sons impulsifs par seconde), le spectre du son répété, l'énergie répandue dans les bandes critiques et, enfin, la sonie, sont des grandeurs qui sont liées entre elles.

5.2.2 Notion d'irrégularité

L'irrégularité du bruit au ralenti diesel est étudiée en vue du développement d'une métrique perceptive liée à la qualité sonore. Les considérations faites sont les suivantes [67].

Considération En tant qu'irrégularités sont considérées les variations brusques de timbre dues à la signature sonore différente des cylindres. L'irrégularité est différente de l'impulsivité du bruit, car elle n'est pas forcément associée aux variations temporelles de l'enveloppe. Les irrégularités dans les cylindres semblent se manifester dans les hautes fréquences (à partir de 1 kHz), alors qu'au-dessous de cette limite, le bruit est jugé plutôt régulier. Le bruit est vu comme la somme de deux types de bruit : le bruit global régulier et le bruit contenant l'irrégularité.

Niveau global et irrégularités L'hypothèse générale est relative à l'équilibre du niveau sonore et le timbre des irrégularités par rapport aux bruit global. Les résultats perceptifs montrent que le fait d'augmenter le niveau sonore des irrégularités rend les bruits perçus plus irréguliers. L'expérience inverse

tendant de montrer la présence des phénomènes de masquage des irrégularités par une augmentation du niveau sonore du bruit global n'est pas confirmé. En revanche, l'observation montre que pour un ralenti variant entre 700 et 1100 tr/min, des bruits à un régime bas sont perçus nettement plus irréguliers qu'à un régime plus élevé. Ceci est expliqué par un masquage temporel qui se fait à haut régime. En conclusion, quatre paramètres influent sur l'irrégularité du bruit sont désignés : le niveau sonore des irrégularités, le niveau sonore du bruit global, le nombre d'irrégularités et le régime moteur.

Impulsivité périodique L'évaluation de la qualité des moteurs diesel comprend le paramètre d'impulsivité périodique. Une étude qui utilise l'analyse de modulation de bande étroite mentionne trois facteurs influant sur la perception de l'impulsivité des bruits diesel [68]. Le premier facteur est relatif aux explosions mêmes et à l'harmonique F2 pour un moteur 4 cylindres. Le deuxième facteur reflète les variations entre les cylindres, car chaque cylindre possède une période propre d'explosion variant de F0.5 et F1. Cette variation équivaut à 7.5 - 15 Hz pour un ralenti à 900 tr/min, l'harmonique F2 étant à 30 Hz. Le troisième facteur est la différence cycle-à-cycle entre les explosions successives de chaque cylindre.

5.2.3 Exemple de configuration rythmique

Le cycle moteur est constitué de 4 explosions différentes. Cet exemple utilise des bruits obtenus avec seulement 2 explosions différentes. Les différentes configurations rythmiques permettent de se faire une idée sur l'impulsivité, la régularité et la périodicité du bruit moteur au ralenti.

Configuration de base La configuration rythmique de base, qu'on va noter "1234", est obtenue à partir de la répétition d'un cycle moteur à 4 explosions différentes (**Fig. 5.7, Piste 43**).

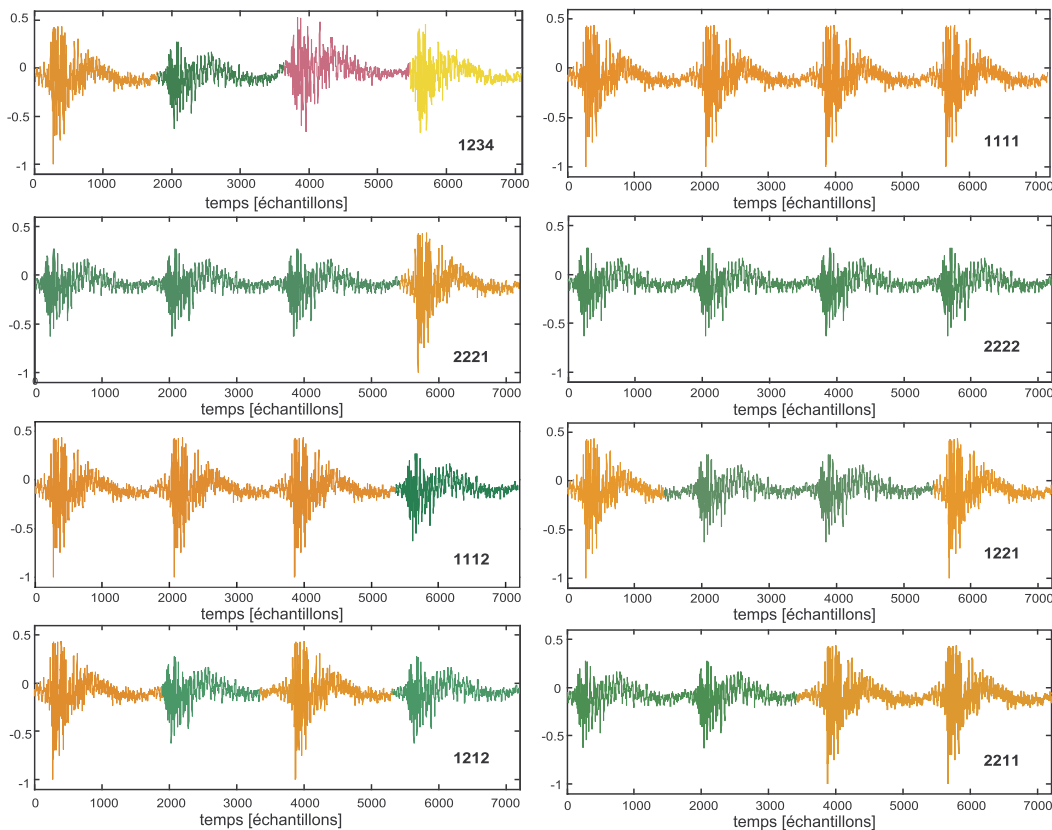


FIG. 5.7 – Configurations issues de deux explosions différentes, signaux temporels

Configuration mono-cylindre A partir du cycle de base "1234", en prenant les deux premiers explosions, on fabrique deux signaux : le premier est constitué de quatre explosions identiques à la première explosion (configuration "1111") ; le deuxième est constitué de quatre explosions identiques à la deuxième explosion (configuration "2222"). Les configurations "1111" (**Piste 44**) et "2222" (**Piste 45**) expriment le bruit d'un monocylindre et son aspect très régulier et impulsif, similaire à un train d'impulsions. Dans ce cas, on

perd la notion d'un cycle constitué de 4 explosions différentes et on perçoit une nouvelle périodicité, déterminée sur un quart de cycle moteur (**Fig.5.8**, période de "1111" et "2222"). Les périodes sont extraites à l'aide des fonctions d'autocorrélation des signaux respectifs.

Une explosion différente Le bruit du monocylindre est un bruit "régulier", puisqu'il est constitué de quatre explosions identiques. Afin de "casser" cette régularité, on introduit une explosion de type "2" dans la configuration du monocylindre "1111" et à l'inverse, une explosion de type "1" dans la configuration du monocylindre "2222". On obtient les configurations "1112" et "2221", respectivement sur **Piste 46** et **Piste 47**. Les deux bruits ainsi obtenus sont moins réguliers que le bruit des configurations monocylindre du fait que les explosions "1" et "2" ont des timbres et des niveaux différents. L'explosion "1" possède une amplitude plus importante que "2" et il semble que ce fait reflète les périodes respectives des configurations "2221" et "1112" (**Fig.5.8**). Dans la configuration "2221", la présence de la plus forte explosion "1" détermine une périodicité égale à un cycle. En revanche, dans la configuration "1112", l'explosion plus faible "2" n'est pas prise en compte et la périodicité extraite reste proche de celle du monocylindre "1111".

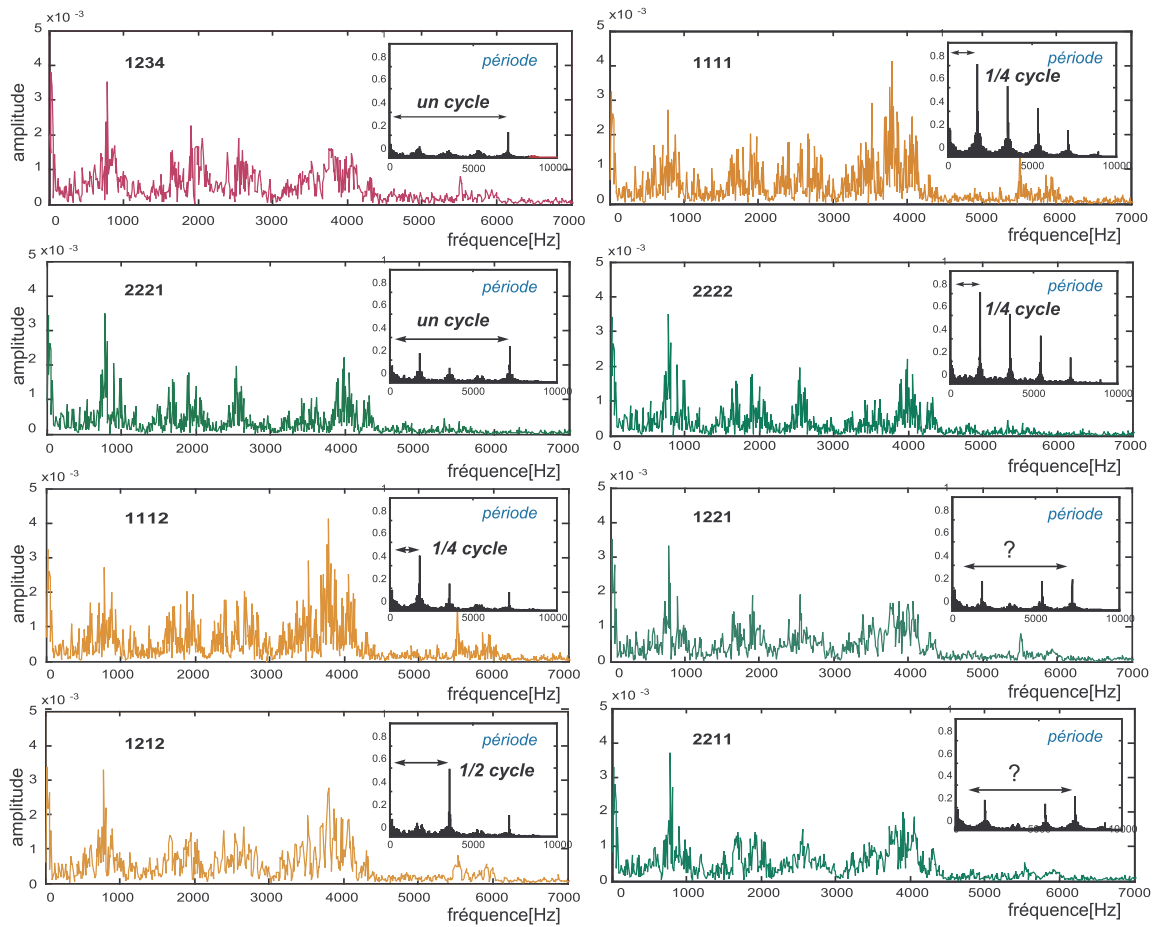


FIG. 5.8 – Spectres et périodes respectifs des configurations de Fig.5.7

Deux explosions différentes En continuant dans la même logique, on remplace cette fois-ci deux explosions dans la configuration du monocylindre. Les nouvelles configurations, "1221" et "2211" (**Piste 48** et **Piste 49**), sont en effet identiques, car lors de la répétition cycliques, les mêmes motifs rythmiques "22" et "11" vont se succéder. Néanmoins, avant de les percevoir comme identiques il est nécessaire un temps d'adaptation de l'ordre de quelques cycles, pendant lesquels la notion du rythme n'est pas tout à fait établie. Ce temps d'adaptation est surtout notable sur la configuration "2211" (**Piste 49**), car elle commence avec les deux explosions "22", plus faibles que "11". Cette observation laisse supposer que l'oreille a la tendance de "se caler" par rapport au temps plus fort. Les périodes extraites pour "1221" et "2211" ne sont pas bien déterminées (**Fig.5.8**). Une configuration qui alterne les explosions "1" et "2" est "1212" (**Piste 50**). Elle possède une période en un demi cycle (**Fig.5.8**). Globalement, en comparant les bruits "1234", "1212" et "1111" possédant une périodicité respectivement un cycle, un demi-cycle et un quart, on perçoit une rapidité croissante dans les événements sonores (**Piste 43**, **Piste 50** et **Piste 44**).

5.2.4 Approche temps-fréquence du bruit au ralenti

Les exemples de configurations mettent en évidence quelques aspects spécifiques des bruits moteurs au ralenti. Ces aspects m'ont incité à proposer une nouvelle approche pour l'étude perceptive des bruits moteur au ralenti.

Particularités des bruits moteur au ralenti Les particularités des bruits moteur au ralenti concernent le timbre, le niveau sonore et la configuration rythmique des explosions. En premier lieu, comme l'a démontré l'exemple ci-dessus, quatre explosions différentes sont nécessaires pour donner la sonorité typique du bruit moteur au ralenti. La présence de moins d'explosions différentes rend le bruit plus régulier, proche d'un train d'impulsions. Au contraire, on peut imaginer que la présence d'un grand nombre d'explosions différentes rendra le bruit aléatoire, sans une période nettement perçue. En deuxième lieu, chaque explosion possède un timbre et niveau sonore différents. Le timbre est grandement influencé par la répartition spectrale. Les aspects temporel et spectral des bruits moteur sont donc étroitement liés. Les différences de niveau sonore et de timbre entre les explosions déterminent le rythme du bruit au ralenti. En troisième lieu, le rythme est déterminé, par définition, comme une succession de temps accentués et non accentués. Dans cette vision, les différences de niveau entre les explosions plus fortes et plus faibles déterminent une séquence rythmique spécifique. Cet aspect rythmique peut changer la perception de la période, en rendant la succession d'explosions plus ou moins rapide.

Nouvelle approche temps-fréquence La nouvelle approche temps-fréquence que je propose est fondée sur les aspects spécifiques du bruit moteur au ralenti. L'objectif est de pouvoir manipuler le niveau sonore et le timbre des explosions, afin d'observer les effets rythmiques sur la perception. Notre outil sera l'analyse temps-fréquence (**Fig. 5.9**) et la synthèse granulaire, présentées en détail dans la deuxième partie de la thèse.

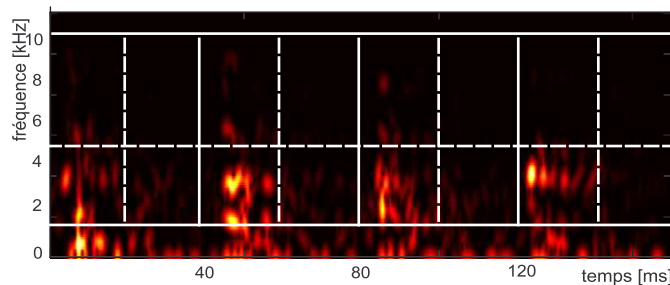


FIG. 5.9 – Spectrogramme temps-fréquence d'un cycle moteur au ralenti et un découpage en zones temps-fréquence

Notre approche se propose de schématiser le bruit moteur en "découpant" le spectrogramme d'un cycle moteur au ralenti en rectangles temps-fréquence. Chaque rectangle représente une zone temps-fréquence qui reflète une partie du timbre des explosions. En apportant des modifications sur les atomes temps-fréquence constituant dans chaque zone, on démontrera les effets perceptifs conséquents sur les bruits de synthèse. Cette expérience fait l'objet de la troisième partie de la thèse.

Conclusion sur le bruit de diesel au ralenti

Le bruit de moteur diesel au ralenti représente une classe de signaux à part. Cette particularité est due à deux raisons principales : le processus de combustion des moteurs diesel et la perception auditive du ralenti. La combustion diesel, provoquée par l'autoinflammation du gazole, sollicite beaucoup plus fortement la structure du moteur diesel que la combustion du moteur à essence. La combustion est la raison principale de l'impulsivité très exprimée du bruit des moteurs diesel. Les facteurs qui influencent la perception du ralenti sont les différences entre les explosions moteur cylindre-à-cylindre et cycle-à-cycle, ainsi que leur contenu temporel et spectral. Le contenu temporel va déterminer plutôt l'aspect impulsif du bruit, alors que le contenu spectral sera responsable du timbre, la signature sonore spécifique de chaque explosion. La nécessité d'étudier les bruits au ralenti à la fois en temps et en fréquence conduit vers l'outil d'analyse le plus adapté : les méthodes temps-fréquence.

Orientation de la thèse : Le choix du ralenti moteur comme l'objet de cette thèse s'est imposé après une large expérience avec des bruits moteurs. La méthode d'analyse-synthèse, développée dans la thèse et présentée dans la deuxième partie du manuscrit, s'est avérée plus adaptée pour le traitement des bruits au ralenti, que pour les bruits moteur à un régime élevé. Ce fait coïncidait favorablement avec un manque relatif de méthodes de synthèse pour les bruits moteur au ralenti. Par ailleurs, les projets de recherche de l'entreprise sur le ralenti diesel au banc ne venaient qu'appuyer ce choix. La nouvelle approche que je propose dans l'étude perceptive des bruits moteur au ralenti peut donner, d'une manière plus globale, des pistes d'exploration des bruits impulsifs répétitifs.

Deuxième partie

Analyse et synthèse temps-fréquence

Résumé de la deuxième partie

La deuxième partie contient l'apport principal de la thèse dans le domaine de traitement du signal. J'ai organisé la structure de cette partie en 5 chapitres, dont quatre discutent la démarche globale d'analyse-synthèse. Le cinquième chapitre présente l'algorithme d'analyse-synthèse développé dans la thèse et son application pour les bruits moteur au ralenti.

Le chapitre 1 introduit les représentations temps-fréquence du signal pour l'analyse du signal simultanément sur les deux dimensions : le temps et la fréquence. L'accent est mis sur les représentations linéaires (Transformée de Gabor, TCFT, ondelettes, échelle perceptive), en vue de la possibilité de synthèse sonore. Les distributions bilinéaires (Wigner-Ville) et leurs versions lissées sont présentées d'une manière informative, car elles sont souvent adoptées dans l'analyse des bruits moteur. Cependant, ces distributions ne sont pas explorées dans la thèse, car elles sont plus difficilement utilisables pour la synthèse des signaux. L'outil d'analyse choisi dans la thèse est la transformée de Gabor.

Le chapitre 2 présente la méthode de réallocation qui se présente comme un post-traitement des images temps-fréquence des signaux. L'objectif principal de la réallocation est la concentration (le déplacement) de l'énergie du signal sur les points temps-fréquence les plus représentatifs. Ce déplacement se fait à l'aide des opérateurs de réallocation qui agissent à la fois en temps et en fréquence. Le résultat est une image plus nette du spectrogramme et la suppression des termes d'interférences pour les distributions bilinéaires. La réallocation est une base sur laquelle repose l'algorithme d'analyse-synthèse proposé dans la thèse.

Le chapitre 3 discute des décompositions atomiques qui présentent le signal projeté sur une famille de fonctions élémentaires, appelées des atomes temps-fréquence. Ici on explore les décompositions parcimonieuses qui comportent un nombre restreint d'atomes pour la reconstruction du signal. Par ailleurs, les méthodes de type "pursuit" ont marqué la recherche dans ce domaine. Je présente ici brièvement leur évolution et les idées communes avec la nouvelle méthode d'analyse-synthèse développée. Cette nouvelle méthode est fondée sur des formules analytiques issues de la Transformée de Gabor, plutôt que sur l'estimation itérative des atomes, comme le fait le "matching pursuit".

Le chapitre 4 expose la synthèse sonore des signaux, en mettant l'accent sur 3 méthodes de synthèse : la synthèse additive, la synthèse par modèles physiques et la synthèse granulaire. C'est la synthèse granulaire qui est adoptée par la suite. Elle porte la même idée que les décompositions atomiques et permet le contrôle en temps et en fréquence de chaque atome. Par ailleurs, les recherches en perception dans le contexte de la synthèse granulaire apportent un lien important avec les tests perceptifs, présentés dans la troisième partie de la thèse.

Le chapitre 5 inclut la nouvelle méthode d'analyse-synthèse des bruits moteur développée dans la thèse. Cette méthode réunit les domaines des quatre chapitres exposés auparavant. Les éléments sur lesquels repose la méthode sont : la transformée de Gabor, les opérateurs de réallocation, les décompositions parcimonieuses avec des critères de choix des atomes et la synthèse granulaire. Chaque étape de l'algorithme est illustrée par des exemples. De plus, les exemples sonores permettent d'écouter le résultat de l'algorithme d'analyse-synthèse sur sons impulsifs de type bruit moteur et de donner un aperçu sur les limites de la méthode. Enfin, suite à des expériences avec plusieurs sons de synthèse, certains phénomènes des sons impulsifs répétitifs sont mis en évidence, afin de servir à la construction des tests perceptifs.

Les représentations temps-fréquence

Les signaux sonores portent de l'information. Dans l'objectif d'extraire cette information, on se sert couramment de deux représentations unidimensionnelles du signal. En temps, c'est la représentation temporelle qui nous permet de suivre l'évolution et les caractéristiques du signal comme la durée, la présence de transitoires ou les changements d'amplitude. En fréquence, à l'aide de l'analyse de Fourier, on dispose du contenu spectral du signal qui nous indique l'aspect harmonique des sons ou, plus globalement, leur timbre. Cependant, ces deux visions se concentrent sur une seule dimension du signal, soit le temps, soit la fréquence. Le spectre d'amplitude de Fourier témoigne du contenu fréquentiel, mais n'indique pas à quel instant en temps les fréquences sont apparues. Or, le spectre de la plupart des signaux qui nous entourent, évolue avec le temps.

Les signaux dont le spectre évolue avec le temps sont appelés non-stationnaires. Ils comportent souvent à la fois des parties tonales et transitoires. Par exemple, les sons des instruments de musiques sont souvent non-stationnaires. Ils débutent par une partie transitoire due à l'attaque de la note, poursuivie par une partie tonale, plutôt stationnaire, et ils finissent par un amortissement global du son. Le besoin d'étudier les sons non-stationnaires fait naître l'analyse conjointe sur les deux dimensions du signal, le temps et la fréquence. Les représentations temps-fréquence expriment davantage la réalité des signaux sonores, car le temps et la fréquence sont liés d'une manière intrinsèque.

6.1 Quelques définitions

Avant de se concentrer sur les représentations temps-fréquence, je vais rappeler quelques définitions essentielles comme la série de Fourier, la transformée de Fourier, le signal analytique, la fréquence instantanée et le retard du groupe.

Séries de Fourier Au début du XIX-ème siècle, Fourier généralise une idée, apparue dans les travaux de Bernoulli et D'Alembert sur la corde vibrante, afin de résoudre l'équation de la chaleur. Cette idée postule que tout mouvement périodique peut être décomposé en simples ondes sinusoïdales. En d'autres termes, tout signal périodique peut s'exprimer sous la forme d'une série trigonométrique, qu'on désigne aujourd'hui comme les séries de Fourier

$$s(t) = a_0 + \sum_{n=1}^{+\infty} (a_n \cos n\omega t + b_n \sin n\omega t). \quad (6.1)$$

Le mouvement périodique possède une fréquence angulaire $\omega_n = \frac{2\pi}{T}n$ qui tient compte de la période, notée T . La fréquence ω_1 représente la pulsation fondamentale et tous les autres termes de la série de Fourier sont obtenus comme des multiples de la fondamentale. Cette caractéristique est propre aux signaux harmoniques.

Transformée de Fourier Toutefois, les séries de Fourier permettent d'explorer seulement l'univers des sons harmoniques, i.e qui ont un caractère périodique. Pour les signaux non-périodiques, l'extension des séries de Fourier est donnée par la transformée de Fourier, qui permet d'associer au signal temporel un spectre continu de fréquences. Le spectre $S(\omega)$ est une fonction de la fréquence ω ,

$$S(\omega) = \int_{-\infty}^{\infty} s(t) e^{-i\omega t} dt \quad (6.2)$$

La transformation inverse de Fourier permet de présenter un signal temporel $s(t)$

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{i\omega t} d\omega \quad (6.3)$$

comme une somme infinie de sinusoïdes $e^{i\omega t}$. Le spectre de Fourier fournit le contenu fréquentiel global du signal, sans préciser à quel instant est apparue chaque fréquence. Cette information peut être éventuellement retrouvée dans la phase.

Signal analytique Une vision plus évolutive du signal est présentée à travers la notion de la fréquence instantanée et du retard du groupe. Afin de définir ces caractéristiques du signal, il est utile de présenter les signaux réels sous une forme complexe d'argument et de phase. L'équivalent complexe d'un signal réel $s(t)$ s'appelle signal analytique $s_a(t)$,

$$s_a(t) = s(t) + iH[s(t)] \quad (6.4)$$

où la partie réelle du signal analytique correspond au signal réel $s(t)$. La partie imaginaire est obtenue à l'aide de la transformée de Hilbert [69]

$$H[s(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(t')}{t - t'} dt' \quad (6.5)$$

Une des particularités du signal analytique réside dans le fait que son spectre \hat{S}_a ne comporte pas de fréquences négatives. En effet,

$$\hat{S}_a(\omega) = \begin{cases} 2\hat{S}(\omega) & , \text{ si } \omega < 0, \\ \hat{S}(0) & , \text{ si } \omega = 0, \\ 0 & , \text{ si } \omega > 0. \end{cases}$$

D'une manière équivalente, on écrit le signal analytique $s_a(t) = A(t)e^{i\phi(t)}$ en termes de module $A(t) = |s_a(t)|$ et de phase $\phi(t) = \arg(s_a(t))$.

Fréquence instantanée et retard de groupe La fréquence instantanée d'un signal porte, comme son nom l'indique, l'information relative à la fréquence présente à un instant donné. Formellement, la fréquence instantanée est la dérivée (positive) en temps de la phase du signal analytique $s_a(t)$

$$\omega_i(t) = \frac{d\phi(t)}{dt} \quad (6.6)$$

Le retard du groupe traduit la même idée que la fréquence instantanée, mais cette fois dans le domaine spectral. Il indique à quel instant est présente une fréquence donnée. Si on note le spectre du signal analytique $S_a(\omega)$ en termes d'amplitude et de phase spectrales, $S_a(\omega) = B(\omega)e^{i\psi(\omega)}$, le retard de groupe est la dérivée de la phase spectrale

$$t_g(\omega) = -\frac{d\psi(\omega)}{d\omega} \quad (6.7)$$

La fréquence instantanée et le retard de groupe sont des fonctions d'une seule variable, soit le temps, soit la fréquence. Elles tracent deux courbes représentatives de certaines caractéristiques du signal dans les repères temps-fréquence - plutôt "horizontale" pour la fréquence instantanée, plutôt "verticale" pour le retard du groupe. Ces courbes ne sont pas réciproques dans le cas général, car elles sont soumises à des propriétés conjointes temps-fréquence du signal. L'exemple du chirp linéaire à enveloppe gaussienne (**Fig. 6.1 a**), confirme le fait que les courbes de la fréquence instantanée et du retard de groupe sont différentes (**Fig. 6.1 b**).

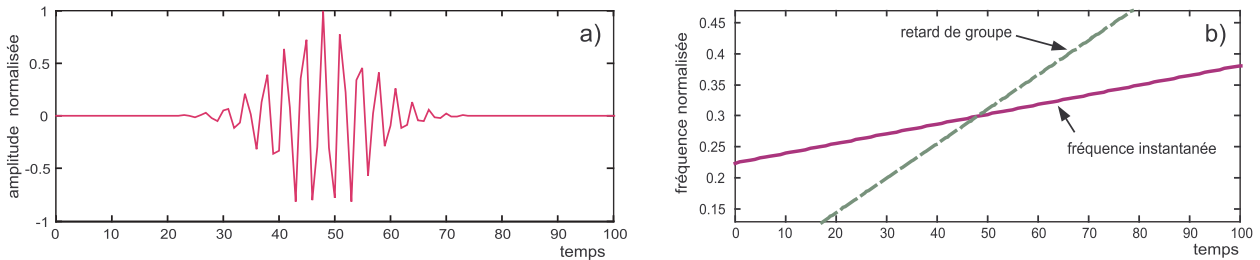


FIG. 6.1 – Un chirp linéaire à enveloppe gaussienne (a) et les courbes de la fréquence instantanée et retard de groupe (b)

La fréquence instantanée et le retard de groupe sont des représentations temps-fréquence unidimensionnelles. Elles sont, en quelque sorte, une transition entre la transformée de Fourier et les représentations temps-fréquence bidimensionnelles, qu'on présente maintenant.

6.2 Transformées de Gabor et de Fourier à court terme

La transformée de Gabor et la transformée de Fourier à court terme (TFCT) font partie des représentations temps-fréquence linéaires. Elles décrivent les changements du spectre à chaque instant de l'évolution temporelle du signal.

6.2.1 Définition

Transformée de Fourier à court terme (TFCT)

La transformée de Fourier à court terme utilise une fonction $h(t)$, appelée fenêtre d'analyse. Cette fenêtre sélectionne une portion du signal et se "déplace" par translation le long de l'axe temporel (**Fig. 6.2**). A chaque translation en temps, on associe le spectre instantané pour la partie du signal qui est analysée par la fenêtre.

Ce processus est appelé "la méthode de la fenêtre glissante" ("moving window method") et consiste à calculer la transformée de Fourier pour chaque portion du signal filtré par la fenêtre $h(t)$. Ainsi est obtenue la définition de la transformée de Fourier à court terme (TCFT) :

$$TFCT(\tau, \alpha) = \int s(t)h^*(t - \tau)e^{-it\alpha}dt. \quad (6.8)$$

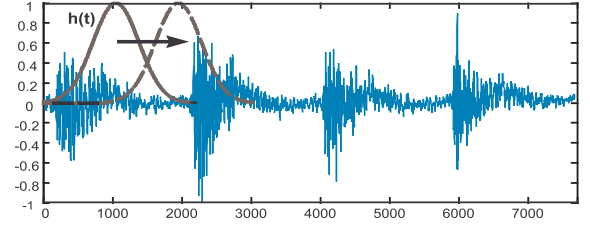


Fig. 6.2 – Filtrage avec la fenêtre d'analyse $h(t)$

Le processus consistant à analyser le signal par petits morceaux découle de l'idée de Gabor en 1946 de décomposer le signal sur une famille de fonctions élémentaires, localisées à la fois en temps et en fréquence [70]. Chaque fonction de la famille de Gabor est obtenue à partir de la fenêtre d'analyse $h(t)$, par translation τ en temps et modulation α en fréquence.

Transformée de Gabor

La transformée de Gabor dans le domaine continu donne une expression de la décomposition du signal sur la famille de fonctions élémentaires

$$G(\tau, \alpha) = \int s(t)h^*(t - \tau)e^{-i(t-\tau)\alpha}dt. \quad (6.9)$$

Cette formule représente aussi le produit scalaire du signal $s(t)$ avec la famille de décomposition $h_{\tau, \alpha}$

$$G(\tau, \alpha) = \langle s(t), h_{\tau, \alpha}(t) \rangle \quad \text{où} \quad h_{\tau, \alpha}(t) = h(t - \tau)e^{i(t-\tau)\alpha}. \quad (6.10)$$

La fenêtre d'analyse $h(t)$ choisie par Gabor

$$h(t) = e^{-\frac{t^2}{2\sigma^2}},$$

est une gaussienne dont le support est déterminé par le paramètre σ . En fait, la gaussienne satisfait les conditions d'une localisation optimale en temps et en fréquence. Cette propriété est illustrée plus loin par le principe d'incertitude de Heisenberg-Gabor.

Spectrogramme

Une forme très intuitive de visualisation est le spectrogramme. Il représente le module au carré de la transformée de Gabor. Le spectrogramme est une fonction à deux variables, le temps et la fréquence (**Fig. 6.3**). Il trace la valeur du spectre de puissance $|G(t_0, \omega_0)|^2$ au point (t_0, ω_0) du plan temps-fréquence. Les spectrogrammes de la TFCT et de Gabor sont équivalents. En revanche, les deux transformées diffèrent en phase. Pour la TCFT, c'est seulement la fenêtre qui est translatée, alors que dans le cas de Gabor, en plus de la fenêtre, c'est également l'exponentielle complexe $e^{it\alpha}$ qui est translatée.

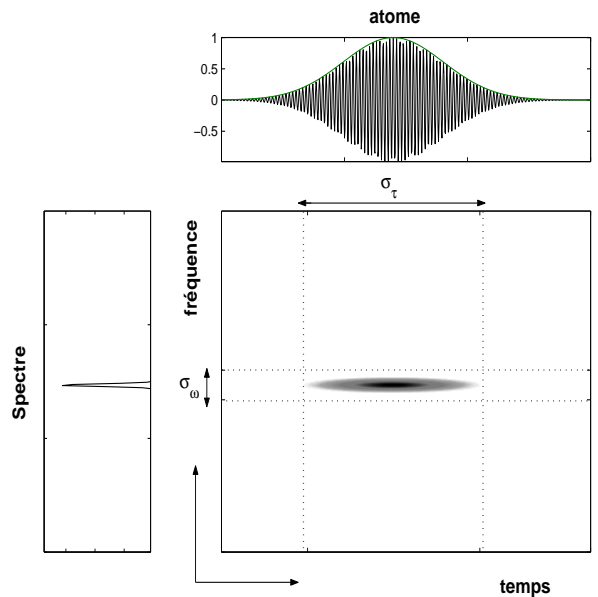


Fig. 6.3 – Un atome temps-fréquence : le signal temporel, le spectre et le spectrogramme

Par défaut, la fenêtre d'analyse de Gabor est la gaussienne. Pour la TFCT, des fenêtres comme Hamming, Hanning, etc., sont souvent utilisées. Dans la littérature de traitement du signal, il n'existe pas de différence rigoureuse entre les deux transformées et elles sont souvent jugées identiques. Dans ce document, on considère principalement la transformée de Gabor à fenêtre gaussienne.

Formule de reconstruction La transformée de Gabor est inversible. Cette propriété permet de revenir au signal du départ $s(t)$ à l'aide de la formule de reconstruction

$$s(t) = \frac{1}{E_h} \int \int G(\tau, \alpha) h(t - \tau) e^{i(t-\tau)\alpha} d\tau d\alpha \quad (6.11)$$

La constante $E_h = 2\pi \int |\hat{h}(\omega)|^2 d\omega$ est déterminée en fonction de l'énergie spectrale de la fenêtre d'analyse. Il est également possible de choisir pour la reconstruction du signal une fenêtre de synthèse $g(t)$, différente de la fenêtre d'analyse $h(t)$, telle que

$$s(t) = \frac{1}{E_{gh}} \int \int G(\tau, \alpha) g(t - \tau) e^{i(t-\tau)\alpha} d\tau d\alpha, \quad (6.12)$$

avec la constante $E_{gh} = 2\pi \int g(\hat{\omega}) \bar{h}(\omega) d\omega$.

Conservation d'énergie La condition d'inversibilité de la transformée de Gabor exige que la fenêtre $h(t)$ soit d'énergie finie, i.e $\int |h(t)|^2 dt < \infty$. L'énergie totale du signal E doit être indépendante du domaine, temporel ou spectral, dans lequel on la calcule. La propriété

$$E = \int |s(t)|^2 dt = \frac{1}{2\pi} \int |S(\omega)|^2 d\omega, \quad (6.13)$$

est connue comme l'égalité de Parseval. La condition de conservation de l'énergie, satisfaite pour les coefficients de Gabor,

$$\int |s(t)|^2 dt = \frac{1}{E_h} \int \int |G(\tau, \alpha)|^2 d\tau d\alpha \quad (6.14)$$

montre que toute l'énergie du signal est contenue dans la représentation $G(\tau, \alpha)$.

6.2.2 Fréquence locale et retard local

D'une manière analogue à la définition de la fréquence instantanée et du retard de groupe d'un signal, on peut déduire des grandeurs similaires, associées à la transformée de Gabor. La forme complexe de la transformée de Gabor, en termes de module $M(\tau, \alpha)$ et de phase $\Phi(\tau, \alpha)$,

$$G(\tau, \alpha) = M(\tau, \alpha) e^{i\Phi(\tau, \alpha)}, \quad (6.15)$$

permet de définir la fréquence locale et le retard local, respectivement comme [71]

$$\frac{d\Phi(\tau, \alpha)}{d\tau} \quad \text{et} \quad \frac{d\Phi(\tau, \alpha)}{d\alpha}. \quad (6.16)$$

Ces grandeurs sont introduites dans le but d'estimer la fréquence instantanée et le retard de groupe des signaux. Deux cas particuliers, le signal monochromatique et l'impulsion de Dirac, mettent en évidence l'estimation directe des paramètres du signal à partir de la fréquence locale et du retard local.

Signal monochromatique Le signal monochromatique $s(t) = e^{i\omega_1 t}$ comporte une seule fréquence ω_1 . Sa transformée de Gabor $G(\tau, \alpha) = \hat{h}(\omega_1 - \alpha) e^{i\omega_1 \tau}$ permet l'estimation directe de la fréquence ω_1 à l'aide de la fréquence locale. Pour un signal monochromatique, la fréquence locale $\frac{\partial \Phi(\tau, \alpha)}{\partial \tau}$ est égale pour tout α à la fréquence recherchée ω_1 . La fréquence du signal ω_1 est donnée par cette fréquence α_1 de la transformée qui représente la solution de l'équation

$$\frac{\partial \Phi(\tau, \alpha)}{\partial \tau} = \alpha \quad (6.17)$$

Par ailleurs, le module de la transformée, $M(\tau, \alpha)$, a une valeur maximale en $\alpha = \omega_1$ pour tout τ . Ceci provient du fait que la fenêtre $h(t)$ est centrée autour du point 0. Or, en ce même point, $\hat{h}(\omega)$ atteint sa valeur maximale (**Fig.6.4 a**). L'estimation de la fréquence du signal peut se faire de deux manières,

soit à partir du maximum de module à un τ fixé, soit à partir de la dérivée de la phase à un α fixé.

Impulsion de Dirac L'impulsion de Dirac $s(t) = \delta(t - t_0)$ qui apparaît à un instant t_0 , possède une transformée de Gabor $G(\tau, \alpha) = \bar{h}(t_0 - \tau)e^{i\alpha(\tau - t_0)}$. Maintenant, l'instant d'apparition t_0 est estimé à partir du retard du groupe $\frac{\partial \Phi(\tau, \alpha)}{\partial \alpha} = \tau - t_0$. Il représente la solution de l'équation

$$\frac{\partial \Phi(\tau, \alpha)}{\partial \alpha} = 0 \quad (6.18)$$

qui est satisfaite en $\tau = t_0$ comme le montre la **Fig. 6.4 b**.

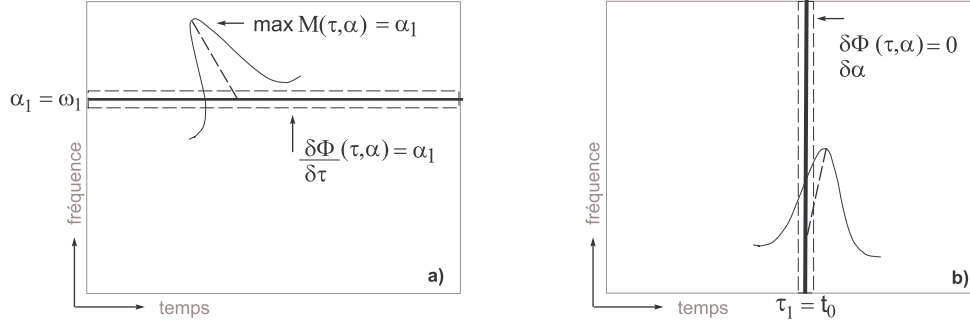


FIG. 6.4 – Extraction des paramètres du signal à l'aide de la fréquence locale (a) et le retard local (b)

Ces propriétés de la transformée de Gabor sont largement exploitées dans le but d'estimer les lois de modulation de fréquence des signaux asymptotiques [72], ainsi que pour la détection des lignes spectrales d'un signal en vue de la synthèse additive [73]. Enfin, les dérivées de la phase permettent de définir des trajectoires caractéristiques du signal dans le plan temps-fréquence (voir par exemple la méthode dite de "arête et squelette" [71]). Ces exemples montrent que certains paramètres du signal peuvent être extraits directement en calculant les dérivées de la phase de la transformée de Gabor. On verra par la suite que la fréquence locale et le retard local sont étroitement liés aux opérateurs de la réallocation qui permettent de concentrer l'information du signal dans le plan temps-fréquence.

6.2.3 Concept atomique

Chaque fonction élémentaire de Gabor est destinée à représenter une information locale du signal, concentrée à la fois en temps et en fréquence. L'ensemble des coefficients associés à ces fonctions détermine entièrement le signal. Par association avec la physique quantique, Gabor appelle la version discrète de la fonction élémentaire, "un quantum de son" (ou encore "logon"), en lui attribuant le sens de la plus petite unité qui constitue le signal [70]. Cette idée s'inscrit dans un large thème philosophique, soulevé par Démocrite, qui pensait que toute matière, peu importe sa nature, était constituée d'atomes. Les atomes étaient des particules si petites que la matière paraissait alors continue [74]. Il n'est pas étonnant alors que la fonction élémentaire de Gabor soit également désignée comme "un atome temps-fréquence".

Aspect physiologique Le concept d'atome a une signification physiologique très importante, car il attribue une nature conjointe temps-fréquence à la perception humaine. Gabor affirme que l'analyse de Fourier du signal sous la forme d'une somme de sinusoides qui ont un support infini en temps, est une vision mathématique correcte, mais loin de la réalité de nos sensations auditives [75]. En effet, malgré le support infini de la sinusoïde, la perception de sa hauteur tonale dépend de la durée du signal en temps, i.e du nombre d'oscillations perçues. Pour des durées très petites (au-dessous de 10 ms), l'estimation de la hauteur peut être fortement perturbée [76]. Gabor souligne le fait que l'oreille tient compte d'une manière inséparable du temps et de la fréquence. Ainsi, il trouve la définition mathématique qui correspond à cette idée : les atomes temps-fréquence.

L'idée atomique est au cœur des méthodes de décompositions atomiques du signal et demeure à l'origine de la synthèse granulaire. Ces méthodes vont faire l'objet des chapitres suivants. Dans la continuité de cette idée se situe également le modèle d'analyse-synthèse développé pour les bruits moteur.

6.2.4 Principe d'incertitude de Heisenberg-Gabor

Les représentations linéaires ne donnent pas une localisation idéale, à la fois en temps et fréquence, des composantes du signal. Par ailleurs, la localisation parfaite d'un atome consisterait à pouvoir diminuer infiniment son support. Or, ce n'est pas possible à cause de la dualité temps-fréquence.

Dualité temps-fréquence La dualité temps-fréquence s'exprime par le fait que toute diminution du support temporel provoque un élargissement du support fréquentiel du signal, et réciproquement. Une illustration de cette idée passe par les deux signaux dont on vient d'observer les caractéristiques : l'impulsion de Dirac et la sinusoïde. D'un côté, l'impulsion de Dirac est parfaitement localisée en temps, mais son spectre s'étale sur toutes les fréquences. D'un autre côté, la sinusoïde porte une seule fréquence, mais son support est infini en temps. L'atome temps-fréquence exprime un cas "intermédiaire", car il est localisé à la fois en temps et en fréquence et possède un support fini sur ces deux dimensions.

Les étendues du signal en temps et en fréquence sont mesurées généralement par les seconds moments normalisés, définis avec les formules

$$\Delta t^2 = \frac{1}{E} \int t^2 |s(t)|^2 dt \quad \text{et} \quad \Delta \omega^2 = \frac{1}{2\pi E} \int \omega^2 |S(\omega)|^2 d\omega \quad (6.19)$$

Les grandeurs Δt^2 et $\Delta \omega^2$ mesurent la variance, i.e l'écart du signal aux valeurs moyennes en temps et en fréquence. Le facteur de normalisation tient compte de l'énergie du signal E , définie par la formule 6.13.

Inégalité de Heisenberg-Gabor Les supports du signal en temps et en fréquence ne peuvent pas être changés indépendamment, car ils sont reliés par l'inégalité

$$\Delta t \Delta \omega \geq \frac{1}{2} \quad (6.20)$$

La relation (6.20) est appelée inégalité de Heisenberg-Gabor. Elle est empruntée au principe d'incertitude, déduit en 1927 par Heisenberg dans le domaine de la mécanique quantique. Pour les signaux, le principe d'incertitude postule qu'il n'est pas possible de construire un signal dont le support est infiniment petit à la fois en temps et en fréquence. La fonction optimale pour laquelle l'inégalité se transforme en égalité est la gaussienne normalisée, pour laquelle $\Delta t \Delta \omega = \frac{1}{2}$. Cette caractéristique principale justifie le choix de Gabor de l'utiliser comme la fonction de décomposition dans sa représentation atomique du signal.

6.2.5 Echantillonnage du schéma de Gabor

La définition de la transformée de Gabor (6.9) et la formule de reconstruction (6.12) sont relatives au domaine continu. Le cas discret revient à découper le plan temps-fréquence en morceaux rectangulaires (**Fig. 6.5**).

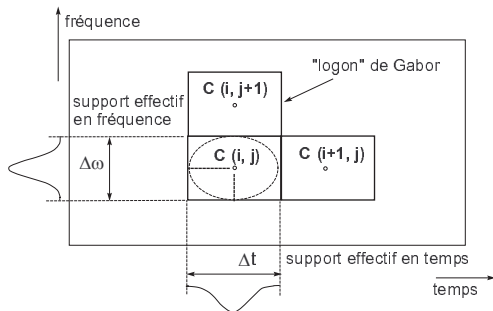


Fig. 6.5 – Le découpage discret de Gabor

Le signal est reconstruit à l'aide de la famille de reconstruction $g_{n,m}(t)$ et des coefficients $G_{[n,m]}$

$$s(t) = \sum_n \sum_m G_{[n,m]} g_{n,m}(t) \quad \text{où} \quad g_{n,m}(t) = g(t - naT) e^{imb\Omega t}. \quad (6.21)$$

La famille de reconstruction est obtenue à partir de la fenêtre de synthèse $g(t)$. Les coefficients de Gabor de la matrice discrète $G_{[n,m]}$ pondèrent chaque élément $g_{n,m}(t)$. C'est justement cet élément que Gabor appelle "logon". Il correspond à un rectangle dans le plan temps-fréquence, centré autour du point $(nt_0, m\omega_0)$. La représentation de Gabor est redondante quand les rectangles se recouvrent entre eux. Ceci

Découpage discret La matrice discrète de Gabor de découpage du plan temps-fréquence a des coordonnées $(\tau = naT, \alpha = mb\Omega)$, où n et m sont des nombres entiers. Deux relations temps-fréquence sont accomplies, $T\Omega = 2\pi$ et $ab \leq 1$. La première est la condition d'échantillonnage de Nyquist pour les signaux, alors que la deuxième est en quelque sorte son équivalence pour l'échantillonnage des fonctions de Gabor [77]. En général, le calcul de la transformée dans tout point introduit une grande redondance. A la place, on calcule plutôt les coefficients de Gabor avec un pas en temps et en fréquence, correspondant aux paramètres de discrétisation a et b .

se manifeste par le fait que la même information du signal est contenue dans deux ou plusieurs éléments adjacents. La redondance semble être une faiblesse de la représentation, mais elle permet, entre autre, les décompositions adaptatives du signal.

Types d'échantillonnage Il existe trois types d'échantillonnage du schéma de Gabor. Dans le cas d'un échantillonnage critique ($ab = 1$), un échantillon du signal analysé est identifié par un seul coefficient de Gabor. Ainsi, il n'y a pas de redondance. Le sur-échantillonnage ($ab < 1$) consiste à avoir plusieurs coefficients de Gabor associés à un échantillon du signal. On définit le degré de sur-échantillonnage avec $Q = \frac{N}{L}$. La grandeur Q mesure combien de fois plus nombreux sont les N coefficients de Gabor, comparés aux L échantillons analysés du signal. En effet, dans le cas d'un sur-échantillonnage infini, la formule de la représentation de Gabor discrète (6.21) tend vers la formule de reconstruction (6.12) dans le domaine continu. Enfin, dans le cas de sous-échantillonnage ($ab > 1$), il existe une perte d'information sur le signal, car le plan temps-fréquence n'est pas entièrement "couvert" avec les fonctions de Gabor.

Echantillonnage critique C'est l'échantillonnage que Gabor adopte initialement. Plus tard, la recherche théorique révèle quelques résultats sur ce type de schéma. Le théorème de Balian-Low, inspiré par le principe d'incertitude, montre que dans le cas d'un échantillonnage critique, il n'est pas possible de construire une base orthogonale avec de bonnes propriétés de localisation [78] [79]. En d'autres termes, la bonne localisation en temps et en fréquence des fonctions de Gabor est incompatible avec l'orthogonalité. Par ailleurs, l'orthogonalité est une propriété mathématique très recherchée, car elle permet la reconstruction du signal $s(t)$ d'une manière unique à partir du produit scalaire entre le signal $s(t)$ et les fonctions de base $g_{n,m}$, i.e

$$s(t) = \sum_m \sum_n \langle s(t), g_{n,m} \rangle g_{n,m} \quad (6.22)$$

tout en préservant l'énergie du signal. Un exemple de décomposition orthogonale est la transformée de Fourier. Cependant, dans le cas de l'échantillonnage critique de Gabor, une base orthonormale peut être obtenue avec un choix approprié de la fenêtre de synthèse, construite biorthogonale à la fenêtre d'analyse [80]. Toutefois, cette solution ne possède pas de bonne localisation. Les solutions proposées souffrent d'une instabilité numérique et il devient nécessaire d'introduire un faible degré de sur-échantillonnage.

Sur-échantillonnage Le cas de sur-échantillonnage est plus propice pour avoir un bon compromis entre les propriétés mathématiques et la capacité de localisation. En revanche, la fenêtre de reconstruction n'est plus unique. Afin de contourner le problème de la base orthogonale, la théorie propose une solution en termes de frame. Selon la définition, un ensemble de fonctions f_k forment une frame, lorsqu'il existe deux constantes A et B pour lesquelles est satisfait

$$A\|s(t)\|^2 \leq \sum_{k \in Z} |\langle s, f_k \rangle|^2 \leq B\|s(t)\|^2. \quad (6.23)$$

Dans le cas de sur-échantillonnage, il est toujours possible de trouver une frame pour une fenêtre gaussienne. Un procédé efficace pour la construction de la fenêtre de synthèse est fourni à l'aide de la transformée de Zak, pour un sur-échantillonnage rationnel [81].

Sous-échantillonnage Le sous-échantillonnage ($ab > 1$) ne permet pas de définir une frame, car la famille de fonctions de Gabor est incomplète. Toutefois, il semble possible de calculer la fenêtre duale [82].

Résumé La discussion précédente met en évidence le fait que, malgré leurs multiples avantages, les fonctions de Gabor (ou encore appelées gaborettes) ne peuvent pas donner une décomposition du signal en termes de base orthogonale appropriée. Cette "faiblesse" a été une des raisons pour chercher d'autres solutions du problème. Le résultats de ces recherches sont les ondelettes.

6.3 Transformée en ondelettes

Les ondelettes sont introduites au début des années 1980 par Morlet [83]. Elles portent la même idée initiale de Gabor de décomposition sur des fonctions élémentaires. La différence est que le plan temps-fréquence est remplacé par la notion du plan temps-échelle. Ainsi, la modulation de la fenêtre d'analyse de Gabor est remplacée par la dilatation de l'ondelette. Les ondelettes trouvent une place incontournable dans les problèmes de traitement des signaux et des images. Elles sont utilisées pour la compression de données, grâce au formalisme appelé multirésolution [84]. Ce procédé utilise les propriétés des filtres et

des bases orthogonales [85]. Ici on va explorer les propriétés d'analyse des ondelettes, plutôt que leur capacité de compression.

6.3.1 Définition et propriétés

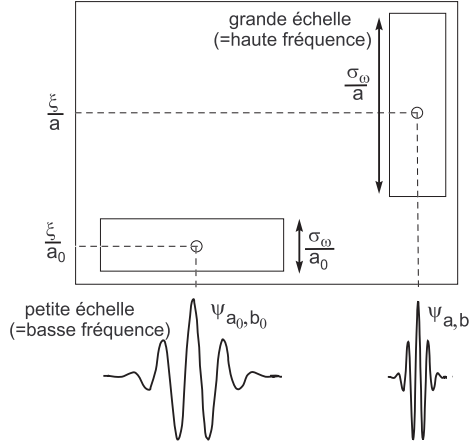
La transformée en ondelettes (TO) est une représentation temps-échelle du signal. Elle transcrit un filtrage du signal avec une fenêtre (ondelette) qui ne reste pas constante pour toutes les fréquences comme c'est le cas de la fenêtre d'analyse de Gabor. L'ondelette adapte son support en fonction du paramètre d'échelle.

Famille d'ondelettes La famille d'ondelettes est obtenue à partir d'une ondelette mère $\psi(t)$ suivant

$$\psi_{(a,b)}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad (6.24)$$

par les opérations de translation en temps à paramètre b et dilatation (ou contraction) en échelle a . À l'échelle a , le paramètre de translation est $\frac{b}{a}$. La transformée en ondelettes est définie comme le produit scalaire du signal $s(t)$ et la famille de fonctions temps-échelle $\psi_{b,a}(t)$,

$$T(b, a) = \int_{-\infty}^{\infty} s(t) \psi_{a,b}^*(t) dt = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(t) \psi^*\left(\frac{t-b}{a}\right) dt. \quad (6.25)$$



Paramètre d'échelle Le paramètre d'échelle a est équivalent au paramètre de fréquence α pour la transformée de Gabor. En fonction de la valeur de a , la fenêtre d'analyse va se contracter ($|a| > 1$) ou se dilater ($|a| < 1$) avec l'augmentation de la fréquence analysée. Dans le premier cas, la fenêtre aura un support large en temps dans les basses fréquences et permettra une meilleure localisation en fréquence. À l'inverse, la contraction de l'ondelette dans les hautes fréquences va privilégier la localisation temporelle, tout en étalant l'information fréquentielle. Le pavage du plan temps-échelle avec les ondelettes est représenté par des rectangles, dont la largeur varie avec l'échelle, contrairement aux gaborettes, où la largeur des rectangles reste constante (**Fig. 6.6**)

Fig. 6.6 – Plan temps-échelle des ondelettes

Relation échelle-fréquence La représentation en échelle semble d'avoir un sens intuitif, car elle est similaire à l'organisation musicale en termes d'octaves. Cependant, en comparaison avec le spectrogramme, où il existe un repère direct par rapport à la fréquence en Hz, la notion de fréquence associée à l'échelle a n'est pas aussi rigoureuse. Certains méthodes de calcul [86] associent l'échelle a à une (pseudo) fréquence f_a en Hz selon la relation

$$f_a = \frac{f_c}{a \cdot \Delta f} \quad (6.26)$$

où Δf est la période d'échantillonnage. La fréquence centrale de l'ondelette, f_c est définie, en termes du barycentre de spectre, comme la fréquence instantanée de l'ondelette au $t = 0$.

Condition d'admissibilité L'ondelette doit satisfaire un certain nombre de critères notamment la condition d'admissibilité

$$C_\psi = 2\pi \int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < +\infty. \quad (6.27)$$

L'existence de l'intégrale (6.27) exige que l'ondelette soit à valeur moyenne nulle, i.e $\int_{-\infty}^{\infty} \psi(t) dt = 0$ et de plus que la transformée en ondelette à la fréquence nulle ($\omega = 0$) soit nulle. De ce fait, $\psi(t)$ est une fonction à support fini en temps, à caractère oscillatoire. C'est une petite onde, d'où le nom ondelette.

En remplissant les conditions, la TO est inversible et on peut revenir au signal original à partir de la formule de reconstruction

$$s(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_0^\infty T(b, a) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) db \frac{da}{a^2}. \quad (6.28)$$

Transformée en ondelettes discrète Similairement à la transformée de Gabor, la transformée en ondelettes possède une version discrète

$$T(m, n) = a_0^{\frac{m}{2}} \int_{-\infty}^{\infty} s(t) \psi(a_0^{-m} t - nb_0) dt \quad (6.29)$$

pour une discrétisation du plan temps-échelle $a = a_0^m$ et $b = b_0^n$. Le cas particulier de $a_0 = 2$ et $b_0 = 1$ désigne la transformée en ondelettes dyadique

$$T(m, n) = \sqrt{2}^{-m} \int_{-\infty}^{\infty} s(t) \psi\left(\frac{t}{2^m} - n\right) dt \quad (6.30)$$

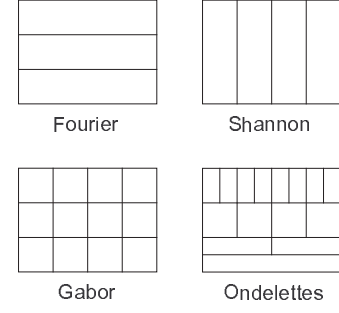


Fig. 6.7 – Le pavage temps-fréquence pour les représentations du signal

Régularité La régularité est une propriété importante des ondelettes, car elle permet de détecter des singularités du signal. Plus précisément, la régularité est assurée par les moments nuls de l'ondelette [87],

$$M_k = \int_{-\infty}^{\infty} t^k \psi(t) dt, \quad (6.31)$$

en sachant que le moment d'ordre 0 est la condition de la moyenne nulle. Les singularités du signal se manifestent sur les coefficients de la TO par une amplitude importante et par une décroissance des valeurs des coefficients avec l'échelle de résolution (**Fig. 6.8**).

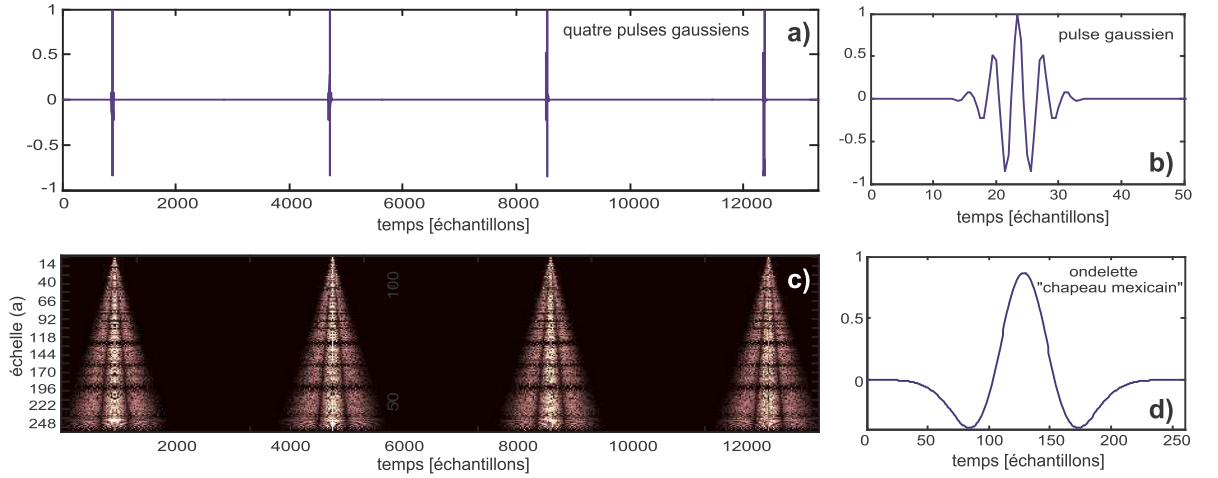


FIG. 6.8 – Quatre atomes gaussiens : signal temporel (a) ; zoom sur un atome (b) ; le scalogramme - la transformée en ondelettes continue (c) avec une ondelette d'analyse de type "chapeau mexicain", obtenue comme la dérivée seconde de la gaussienne (d)

Maxima du module de la TO Les maxima du module de la transformée en ondelettes portent une information significative sur le signal. Les maxima dépendent de l'échelle a de la transformée $T(a, b)$. Le point (a_0, t_0) comporte un extremum local si la première dérivée $\frac{dT(a_0, t)}{dt}$ passe par un zéro au point $t = t_0$. Pour avoir un maximum du module au point (a_0, t_0) , il faut que les valeurs voisines soient plus petites,

$$|T(a_0, t)| < |T(a_0, t_0)| \quad (6.32)$$

i.e lorsque t est soit dans la partie gauche, soit dans la partie droite du voisinage du point t_0 . Lorsque t appartient à une autre partie du voisinage de t_0 , la condition $|T(a_0, t)| \leq |T(a_0, t_0)|$ doit être satisfaite. Une ligne de maxima est une courbe dans l'espace (a, t) qui comporte les maxima locaux du module de la transformée en ondelettes. La définition des maxima pour la transformée de Gabor se fait d'une manière équivalente.

6.3.2 Analyse multirésolution

La transformée en ondelettes, tout comme la transformée de Gabor, peut être interprétée au sens d'un filtrage du signal avec un banc de filtres.

Banc de filtres Une ondelette possède un support fini en temps et son spectre dans le domaine de Fourier s'écrit

$$\hat{s}(at) = \frac{1}{|a|} \hat{s}\left(\frac{\omega}{a}\right) \quad (6.33)$$

Cette formule exprime la dualité temps-fréquence au sens où une dilatation de l'ondelette dans le domaine temporel d'un facteur 2 va compresser la largeur du spectre et le décaler d'un facteur 2. Dans les basses fréquences, le support très large en temps pose un problème pour recouvrir la totalité du domaine spectral, car il demande l'utilisation d'un grand nombre d'ondelettes. En introduisant une fonction d'échelle $\phi(t)$ qui possède un spectre passe-bas et satisfait la condition $\int \phi(t) = 1$, on recouvre tout le spectre (**Fig. 6.9**). Des ondelettes comme la gaussienne, le chapeau mexicain ou l'ondelette de Morlet n'ont pas une fonction d'échelle.

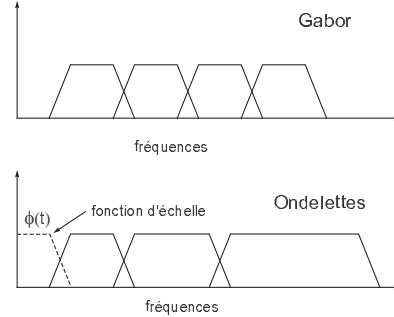


Fig. 6.9 – Interprétation en banc de filtres des transformées de Gabor et en ondelettes

Espace d'approximation Le découpage de l'espace fréquentiel par le banc de filtres permet la définition de l'algorithme de multi-résolution. L'idée consiste à analyser à la fois les niveaux qui comportent une information grossière du signal (les basses fréquences), en complétant l'approximation avec les niveaux plus fins (les hautes fréquences). Les niveaux grossiers sont décrits par l'espace d'approximation, alors que les niveaux fins appartiennent à l'espace de détail. L'espace d'approximation est constitué de sous-espaces V_j , imbriqués les uns dans les autres, et qui sont associés à un changement de facteur d'échelle. Le signal $s(t)$ et le signal $s(\frac{t}{2})$, correspondant à une dilatation de facteur 2, appartiennent respectivement à V_j et V_{j+1} . Le passage de l'espace V_j vers l'espace V_{j+1} correspond à un zoom sur le signal. L'ensemble des sous-espaces $\dots \subset V_1 \subset V_0 \subset V_{-1} \dots V_{j+1} \subset V_j$ recouvrent complètement l'espace du signal.

La fonction d'échelle $\phi(t)$ a la particularité d'engendrer une base orthonormée de V_j par dilatation et translation

$$\phi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \phi(2^{-j}t - n) \quad (6.34)$$

Les fonctions de base $\phi_{j,n}(t)$ permettent la projection du signal $s(t)$ à l'aide du produit scalaire sur les fonctions de base. Ainsi l'approximation à la résolution j est donnée par les coefficients $a_n^j = \langle s, \phi_{j,n} \rangle$.

Espace de détail L'espace de détail W_j complète l'information de l'espace d'approximation, en comportant la différence entre l'approximation à la résolution j et l'approximation à la résolution $j-1$,

$$V_{j-1} = V_j \oplus W_j \quad (6.35)$$

où W_j est le supplémentaire orthogonal de V_j dans V_{j-1} . Les espaces de détail sont orthogonaux entre eux, i.e $W_j \perp W_k$ pour tout $j, k \neq j$. De la même façon que la famille de la fonction d'échelle engendre une base orthonormée dans l'espace des d'approximation V_j , les ondelettes $\psi_{j,n}$

$$\psi_{j,n} = 2^{-\frac{j}{2}} \psi(2^{-j}t - n) \quad (6.36)$$

forment une base orthonormée dans l'espace des détails W_j .

A l'issue de l'algorithme de multi-résolution, le signal $s(t)$ se décompose sur les fonctions de base des espaces d'approximation et de détail suivant

$$s(t) = \sum_{j,n} \langle s, \psi_{j,n} \rangle \psi_{j,n}(t) = \sum_n \langle s, \phi_{j,n} \rangle \phi_{j,n}(t) + \sum_{k \leq j,n} \langle s, \psi_{k,n} \rangle \psi_{k,n}(t). \quad (6.37)$$

L'analyse de multi-résolution est un algorithme remarquable, car il permet plusieurs applications sur les signaux et les images. En pratique les coefficients sont estimés à l'aide des techniques de filtrage numérique [88].

Analyse et reconstruction Les projections orthogonales sur les différents sous-espaces sont implémentées par un banc de filtres. Un banc de filtres sous-échantillonnés convolue le signal avec un filtre passe-bas, représentant l'espace d'approximation, et un filtre passe-haut, représentant l'espace de détail (**Fig. 6.10**). Si on considère $h[n]$ la réponse impulsionnelle d'un filtre numérique H , les coefficients a_n^j peuvent être obtenus d'une manière récursive à partir de

$$a_n^j = \sum_l \tilde{h}[2n - l] a_n^{j-1} \quad (6.38)$$

où $\tilde{h}[n] = h(-n)$ est le filtre symétrique de H . Une procédure similaire permet d'aboutir à l'expression des coefficients de détails

$$d_n^j = \sum_l \tilde{g}[2n - l] a_n^{j-1} \quad (6.39)$$

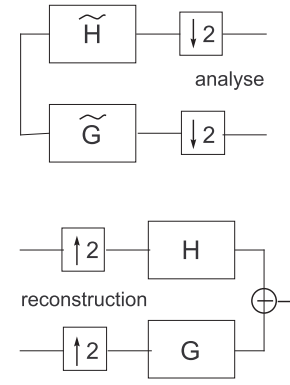


Fig. 6.10 – Filtres d'analyse et synthèse

L'analyse de multi-résolution a permis de diverses applications des ondelettes sur les signaux sonores et les images.

6.3.3 Types d'ondelettes

L'intérêt que les ondelettes ont suscité en vue des applications diverses, a engendré la construction de nombreuses familles d'ondelettes. Une classification distingue 5 types d'ondelettes [86].

Le premier type est le plus simple, et contient trois ondelettes de base : la gaussienne, l'ondelette de Morlet et l'ondelette "chapeau mexicain". Elles sont essentiellement prévues pour une analyse en ondelettes continue. Ces ondelettes n'ont pas de fonction ϕ , sont symétriques et ne sont pas orthogonales. L'avantage est que la fonction ψ possède une formule explicite, mais elle n'a pas de support compact et la reconstruction du signal n'est pas assurée.

Le deuxième groupe d'ondelettes est représenté par les ondelettes de Meyer. Pour ce groupe, la fonction ψ existe et l'analyse est orthogonale. Les ϕ et ψ sont dérivables, mais n'ont pas de support compact. Pareillement, comme pour le premier groupe, elles sont symétriques, mais il n'est pas possible d'introduire un algorithme rapide de calcul.

Le troisième type comprend les ondelettes orthogonales ayant un support compact. Ce sont les ondelettes de Daubechies, les coiffettes, qui ont de très bonnes propriétés comme le support compact et l'analyse orthogonale.

Au quatrième type appartiennent les ondelettes biorthogonales comme les B-splines. Les avantages sont que ψ et ϕ ont un support compact et pour l'analyse, et pour la reconstruction. L'analyse n'est pas orthogonale mais biorthogonale.

Le cinquième type regroupe les ondelettes complexes, comme la gaussienne complexe, la Morlet complexe, la Shannon complexe, etc. Les avantages et les inconvénients de ce groupe ressemblent à ceux du premier groupe, au sens où il n'y a pas d'algorithme rapide, ni d'orthogonalité, en revanche la fonction ψ possède une expression explicite.

6.4 Distribution de Wigner-Ville

La distribution de Wigner (ou plus couramment Wigner-Ville) est définie en 1932 par Wigner dans le domaine de la mécanique quantique et introduite pour le traitement des signaux en 1948 par Ville. C'est une représentation bilinéaire (ou encore dite quadratique) du signal. De ce fait, elle diffère considérablement des représentations linéaires comme telles de Gabor ou des ondelettes.

6.4.1 Définition et propriétés

La distribution de Wigner-Ville d'un signal $s(t)$ représente, pour un instant donné, la transformée de Fourier de la quantité $s^*(t - \frac{1}{2}\tau).s(t + \frac{1}{2}\tau)$, où $s^*(t)$ est le complexe conjugué du signal $s(t)$.

Définition La distribution de Wigner-Ville s'écrit dans le domaine temporel et spectral comme

$$W(t, \omega) = \int s(t + \frac{\tau}{2}).s^*(t - \frac{\tau}{2}) e^{-i\tau\omega} d\tau \quad \text{et} \quad W(t, \omega) = \int S(\omega + \frac{\xi}{2}).S^*(\omega - \frac{\xi}{2}) e^{it\xi} d\xi. \quad (6.40)$$

L'aspect bilinéaire s'exprime par le fait que la formule de la distribution fait intervenir une forme quadratique du signal $s(t)$. Le paramètre τ traduit le décalage en temps.

La distribution de Wigner-Ville (DWV) est toujours réelle, même pour les signaux complexes. En revanche, elle n'est pas strictement positive, mais possède aussi des valeurs négatives. La DWV est inversible, car en inversant

$$s^*(t - \frac{\tau}{2}) \cdot s(t + \frac{\tau}{2}) = \int W(t, \omega) e^{i\tau\omega} d\omega, \quad (6.41)$$

et en posant $t = \tau/2$, $k = 1/s^*(0)$ et $\tau = t$, on obtient la formule de reconstruction du signal $s(t)$,

$$s(t) = k \int W(\frac{t}{2}, \omega) e^{it\omega} d\omega. \quad (6.42)$$

La formule (6.42) montre que le signal peut être reconstruit, à une constante près, à partir de la distribution de Wigner-Ville.

Marginales et covariance La DWV satisfait les conditions des marginales temps-fréquence : les densités marginales sont égales aux densités d'énergie, à la fois dans le domaine temporel et spectral

$$\int W(t, \omega) d\omega = |s(t)|^2, \quad (6.43)$$

$$\int W(t, \omega) dt = |S(\omega)|^2 \quad (6.44)$$

La relation $E = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(t, \omega) dt d\omega$, exprime la conservation de l'énergie du signal. Par ailleurs, les conditions 6.43 et 6.44 ne sont pas satisfaites à la fois pour le spectrogramme.

La covariance par translation en temps et en fréquence est une autre propriété importante de la distribution de Wigner-Ville. Elle s'exprime comme

$$\begin{aligned} y(t) = s(t - t_0) &\Rightarrow W_y(t, \omega) = W_s(t - t_0, \omega) \\ y(t) = s(t) e^{i\omega_0 t} &\Rightarrow W_y(t, \omega) = W_s(t, \omega - \omega_0) \end{aligned}$$

Ces relations sont spécifiques à une classe de distributions, la classe de Cohen, dont la DWV fait partie.

Localisation optimale La distribution de Wigner-Ville assure une localisation optimale des impulsions de dirac dans le domaine temporel, des signaux monochromatiques dans le domaine fréquentiel et des chirps linéaires dans le domaine temps-fréquence [69]. L'énergie du chirp linéaire est parfaitement localisée par la DWV autour de la ligne de la fréquence instantanée $\omega_i = \omega_0 + \beta t$

$$s(t) = e^{i[\omega_0 t + \frac{\beta}{2} t^2]} \Rightarrow W_s(t, \omega) = \delta(\omega - (\omega_0 + \beta t)) \quad (6.45)$$

comme on peut le voir sur la **Fig. 6.11, en bas**. L'exemple de **Fig. 6.11**, réalisé à l'aide de la Time-Frequency Toolbox [89], compare les propriétés de localisation de la DWV et de la transformée de Gabor avec des paramètres d'échantillonnage différents sur un signal test : un chirp linéaire d'une longueur de 512 échantillons. Le spectrogramme de Gabor possède comme paramètres d'échantillonnage : N , le nombre de coefficients de Gabor et Q , le degré de sur-échantillonnage. On observe que la meilleure localisation de la transformée de Gabor, avec des paramètres ($N = 128, Q = 64$) (**Fig. 6.11, en haut**), reflète une moins bonne estimation de la densité d'énergie et à l'inverse pour la transformée de Gabor, avec des paramètres ($N = 256, Q = 256$), (**Fig. 6.11, au milieu**). Le spectrogramme ne peut pas donc fournir

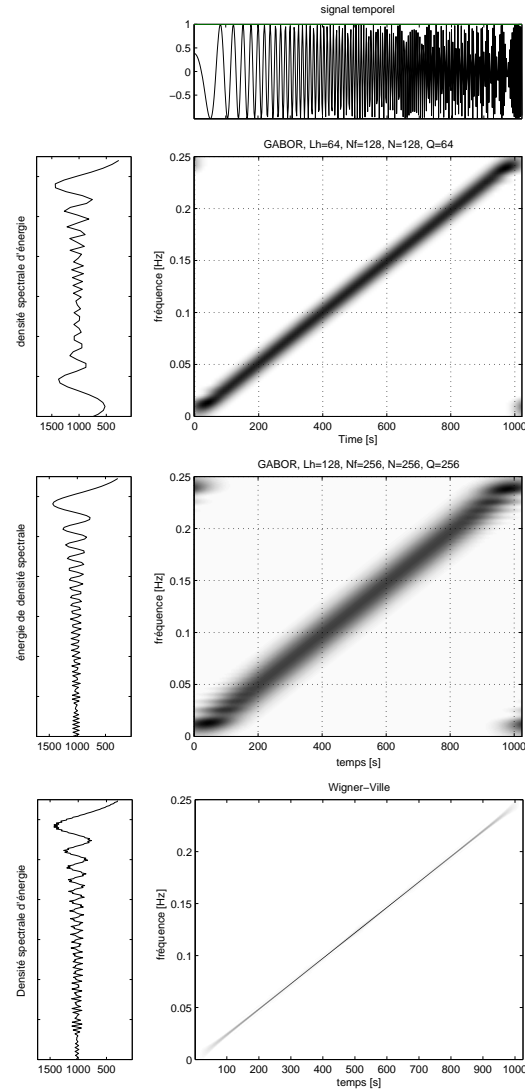


Fig. 6.11 – Comparaison entre la transformée de Gabor et la DWV : localisation d'un chirp linéaire. Gabor ($N = 128, Q = 64$) (en haut), Gabor ($N = 256, Q = 256$) (au milieu) et DWV (en bas). Pour plus de précisions, voir le texte.

une localisation idéale comme la DWV. En revanche, une présentation proche de la DWV est possible à l'aide des méthodes de réallocation, appliquées pour le spectrogramme. Ces méthodes font l'objet du chapitre suivant.

DWV et la fréquence instantanée L'exemple du chirp linéaire est un cas particulier de la relation plus générale et très importante entre la DWV et la fréquence instantanée du signal. La fréquence instantanée ω_i du signal $s(t)$ peut être exprimée comme le moment d'ordre 1 de la distribution de Wigner-Ville comme

$$\omega_i(t) = \frac{\int_{-\infty}^{\infty} \xi W_{s_a}(t, \xi) d\xi}{\int_{-\infty}^{\infty} W_{s_a}(t, \xi) d\xi}, \quad (6.46)$$

où s_a est le signal analytique qui correspond à $s(t)$. La relation (6.46) s'interprète comme un centre de gravité. Une formulation similaire à l'aide de la distribution de Rihaczek est utilisée par Koderà pour la définition des opérateurs de la réallocation [90].

6.4.2 Interférences

La distribution de Wigner-Ville est particulièrement adaptée pour l'analyse des signaux avec une seule composante. En revanche, le traitement des signaux multicomposantes fait apparaître des termes d'interférence. Pour un signal constitué de deux composantes la distribution de Wigner-Ville n'est plus simplement une somme des distributions de chaque composante.

La DWV d'un signal multicomposante La bilinéarité de la DWV fait apparaître les termes croisés dûs aux produits deux par deux des distributions de chaque composante. Ainsi, pour un signal de deux composantes, i.e $s = s_1 + s_2$, la distribution de Wigner-Ville s'écrit

$$W_{s_1+s_2}(t, \omega) = W_{s_1}(t, \omega) + W_{s_2}(t, \omega) + 2\Re \{W_{s_1 s_2}(t, \omega)\} \quad (6.47)$$

et dans cette expression participe la distribution de Wigner-Ville croisée

$$W_{s_1 s_2}(t, \omega) = \int s_1\left(t + \frac{\tau}{2}\right) \cdot s_2^*\left(t - \frac{\tau}{2}\right) e^{-i\tau\omega} d\tau \quad (6.48)$$

qui n'est plus réelle, mais complexe. Les termes croisés proviennent du fait que la DWV tient compte non seulement de la valeur locale du signal à un moment donné, mais également du "passé" et du "futur" du signal. Par ailleurs, la formule 6.47 peut être généralisée pour un signal de plus de deux composantes. Un inconvénient majeur des interférences est qu'elles ne portent aucune signification physique sur la nature du signal. De ce fait, elles rendent ambiguë la lecture d'une représentation temps-fréquence du type Wigner-Ville. En effet, il est difficile de distinguer la partie propre au signal et celle qui apparaît suite à des interférences et qui se superpose aux vraies valeurs des composantes du signal.

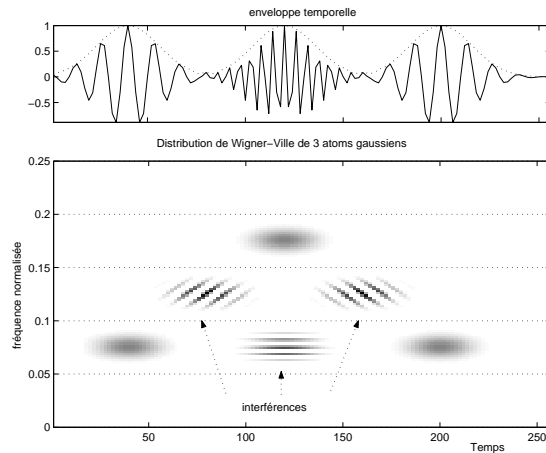


Fig. 6.12 – Trois atomes temps-fréquence et leur distribution de Wigner-Ville. Les termes d'interférence sont présents entre chaque composante

Exemple Cet effet est visible sur la Fig. 6.12. Elle montre trois atomes gaussiens et leur représentation de Wigner-Ville. Les interférences surgissent géométriquement à mi-distance entre chaque composante. La position relative des atomes dans le plan temps-fréquence modifie aussi la période et la direction des oscillations des termes d'interférence [89].

En règle générale, plus le nombre de composantes du signal est élevé, plus la lecture de la distribution devient difficile, car les interférences se superposent aux vraies composantes du signal. De ce fait, la formulation originale de Wigner-Ville n'est pas adaptée aux traitements des signaux multicomposante, dont une grande majorité des signaux, les bruits moteur y compris font partie. En revanche, dans la mesure où les interférences peuvent être correctement interprétées, elles portent une information supplémentaire sur la phase du signal.

Curieusement, la présence des interférences dans la distribution de Wigner-Ville est inévitable, mais aussi nécessaire pour assurer ses bonnes propriétés théoriques, citées précédemment [91].

6.4.3 Généralisation - classe de Cohen

La définition de la distribution de Wigner-Ville exige la connaissance de la quantité

$$s^*(t - \frac{1}{2}\tau).s(t + \frac{1}{2}\tau) \quad (6.49)$$

sur toute la plage temporelle, de $\tau = -\infty$ à $\tau = +\infty$. En réalité, un tel intervalle peut poser un problème de calcul. De plus, très souvent on s'intéresse à l'analyse du signal localement, autour d'un instant t , sans pour autant se soucier des instants éloignés.

Pseudo Wigner-Ville Ce raisonnement fait que la quantité (6.49) est souvent fenêtrée par une fonction $h(\tau)$, centrée autour de $\tau = 0$. Ainsi, on obtient une nouvelle distribution, appelée "pseudo Wigner-Ville",

$$PW(t, \omega) = \int h(\tau) s^*(t - \frac{1}{2}\tau).s(t + \frac{1}{2}\tau) e^{-i\omega\tau} d\tau \quad (6.50)$$

En effet, la version "pseudo" représente un lissage en fréquence de la distribution de Wigner-Ville,

$$PW(t, \omega) = \int H(\omega - \xi) W(t, \xi) d\xi \quad (6.51)$$

où $H(\omega)$ est la transformée de Fourier de $h(t)$. Le résultat du lissage est une diminution des interférences. Par contre, les bonnes propriétés comme les marginales et la fréquence instantanée de la DWV sont perdues.

Classe de Cohen La classe de Cohen résume toutes les représentations temps-fréquence, covariantes par translation en temps et en fréquence

$$C(t, \omega) = \int \int \int e^{i\xi(\eta-t)} s^*(\eta - \frac{1}{2}\tau) s(\eta + \frac{1}{2}\tau) f(\xi, \tau) e^{-i\tau\omega} d\eta d\tau d\xi. \quad (6.52)$$

C'est une formule générale, à partir de laquelle on peut déduire toute représentation temps-fréquence. La fonction de deux variables $f(\xi, \tau)$ s'appelle "noyau" ("kernel" en anglais). La distribution de Wigner-Ville possède un noyau qui vérifie la condition $f(\xi, \tau) = 1$. La formulation plus compacte de (6.52) est la suivante

$$C(t, \omega, \Pi) = \int \int \Pi(\eta - t, \xi - \omega) W(\eta, \xi) d\eta d\xi, \quad (6.53)$$

où $\Pi(t, \omega)$ est la transformée de Fourier bi-dimensionnelle de la fonction $f(\xi, \tau)$,

$$\Pi(t, \omega) = \int \int f(\xi, \tau) e^{-i(\omega\tau + \xi t)} dt d\omega. \quad (6.54)$$

Pour la DWV, la propriété $\Pi(t, \omega) = \delta(t)\delta(\omega)$ explique la localisation parfaite de la sinusoïde et l'impulsion de Dirac. D'une manière générale, il est possible de contrôler le lissage indépendamment en temps et en fréquence, en présentant $\Pi(t, \omega) = g(t)H(-\omega)$ comme une fonction séparable en temps et en fréquence. Ainsi est obtenue la distribution "pseudo Wigner-Ville lissée"

$$SPW(t, \omega) = \int h(\tau) \int g(\eta - t) s^*(\eta - \frac{1}{2}\tau).s(\eta + \frac{1}{2}\tau) e^{-i\omega\tau} d\eta d\tau, \quad (6.55)$$

qui est une des distributions les plus flexibles par rapport à l'ajustement du lissage séparément en temps et en fréquence.

Lien avec le spectrogramme La distribution bilinéaire de Wigner-Ville est reliée aux transformées de Gabor et en ondelettes continues. Les modules au carré des deux transformées linéaires peuvent être interprétés comme des versions lissées de la distribution de Wigner-Ville du signal [91]. Les relations suivantes sont valables pour le spectrogramme de Gabor [92]

$$|G(\tau, \alpha)|^2 = \int W_s(\eta, \xi).W_g(\eta - \tau, \xi - \alpha) d\eta d\xi, \quad (6.56)$$

$$|T(b, a)|^2 = \int W_s(\eta, \xi).W_\psi(\frac{\eta - b}{a}, a\xi) d\eta d\xi \quad (6.57)$$

où W_g et W_ψ sont les distributions de Wigner-Ville des fonctions de lissage respectivement pour le spectrogramme et pour le scalogramme.

DWV ou spectrogramme ? En comparant la distribution de Wigner-Ville et le spectrogramme, on ne peut pas juger, pour tout type de signaux, laquelle des deux représentations est meilleure. D'un côté, il est incontestable que la DWV satisfait les conditions des marginales et conserve les supports. Elle donne ainsi une image très précise du signal, de la fréquence instantanée et du retard de groupe. Cette précision ne peut pas être obtenue directement du spectrogramme. D'un autre côté, le spectrogramme dépend du choix de la fenêtre d'analyse et de l'échantillonnage, alors que ces problèmes ne se posent pas pour la DWV. En même temps, le spectrogramme possède une valeur toujours positive, alors que ce n'est pas toujours le cas pour la DWV, dont les termes d'interférences ont des oscillations négatives. Par ailleurs, les termes d'interférence de la DWV représente une faiblesse de la distribution, et en pratique la version "pure" (i.e non-lissée) de Wigner-Ville est rarement utilisée. Toutefois, même si le lissage améliore les interférences, il fait également perdre à la DWV ses propriétés théoriques. En conclusion, malgré les capacités de localisation de la distribution de Wigner-Ville, il me semble que l'interprétabilité du spectrogramme reste nettement plus abordable.

6.5 Représentation en échelle perceptive

La transformée de Gabor et la TFCT peuvent être vues comme une analyse du signal au travers d'un banc de filtres, où chaque filtre est localisé autour d'une fréquence analysée. Pour Gabor et la TFCT, les filtres ont une largeur constante pour toutes les fréquences du signal. Toutefois, cette convention simple ne représente pas fidèlement la perception auditive de l'oreille humaine. Mieux adaptée à la perception de l'oreille, la transformée en ondelettes se sert d'une fenêtre qui varie en fonction de la fréquence. Plus loin, le besoin d'apporter des connaissances plus précises du filtrage effectué par l'oreille a ouvert le chemin des représentations en échelle perceptive.

Processus auditif dans la cochlée Les représentations en échelle perceptive sont basées sur une modélisation de la perception des sons dans la cochlée. En effet, le comportement de la membrane basilaire ressemble à celui d'un banc de filtres. Les filtres occupent une bande fréquentielle qui est plus étroite dans les basses fréquences et qui devient de plus en plus large dans les hautes fréquences. Toutefois, les filtres auditifs ne sont pas exactement équivalents au banc de filtres de la famille d'ondelettes. Des exemples principales représentations à échelle perceptive sont l'échelle de Bark, les filtres ERB et les filtres Gammatone [93]. En effet, les filtres Gammatone représentent une alternative perceptive à la fenêtre d'analyse des représentations temps-fréquence, l'ondelette des représentations temps-échelle ou le noyau de lissage des distributions bilinéaires de la classe de Cohen.

6.5.1 Echelle de Bark

L'échelle de Bark repose sur le concept des bandes critiques. Ce concept remonte à Fletcher qui suppose en 1940 que le système auditif fonctionne comme un banc de filtres [94]. Fletcher teste cette hypothèse pour le masquage d'un son pur par un bruit blanc. Le masquage se manifeste par le fait que le son devient moins audible dans la présence du bruit, i.e le bruit masque le son. Fletcher constate que seulement une portion du spectre du bruit est nécessaire pour masquer le son. Pour déterminer la largeur de cette portion du spectre, il suppose qu'au seuil du masquage les intensités sont égales. Fletcher appelle cette largeur "bande critique" comme le rapport entre l'intensité du son pur au seuil et la densité spectrale du bruit. Ce mécanisme de filtrage se révèle par la suite valable non seulement pour le masquage, mais également pour d'autres propriétés perceptives.

Bande critique La définition de "bande critique" qu'on connaît actuellement est donnée par Zwicker et Feldtkeller [95]. Cette définition est issue des travaux sur la sonie. Plus précisément, la largeur de la bande critique est déterminée comme la zone où la sonie d'une bande de bruit de niveau reste constante tant que la largeur de la bande ne dépasse pas une certaine limite. Afin d'éviter la confusion, la définition de "bande critique", donnée auparavant par Fletcher est désignée désormais comme "rapport critique". L'interprétation des bandes critiques s'inscrit dans le mécanisme global de filtrage effectué par l'oreille. Le filtrage repose sur le concept que l'oreille ne traite pas d'une manière similaire toutes les fréquences, mais effectue un découpage de la plage audible et traite d'une manière similaire les fréquences à l'intérieur de chaque bande fréquentielle.

Zwicker et Feldtkeller présentent la plage des fréquences de 20 Hz à 16 kHz découpée en 24 bandes

critiques, appelées bandes de Bark ¹. Chaque bande critique possède une fréquence centrale, ainsi que deux fréquences de coupure, inférieure et supérieure, qui déterminent la largeur de la bande (**Fig. 6.13**). Jusqu'à 500 Hz, les bandes ont une largeur constante de 100 Hz. Au-delà de 500 Hz, la largeur croît d'un ordre logarithmique.

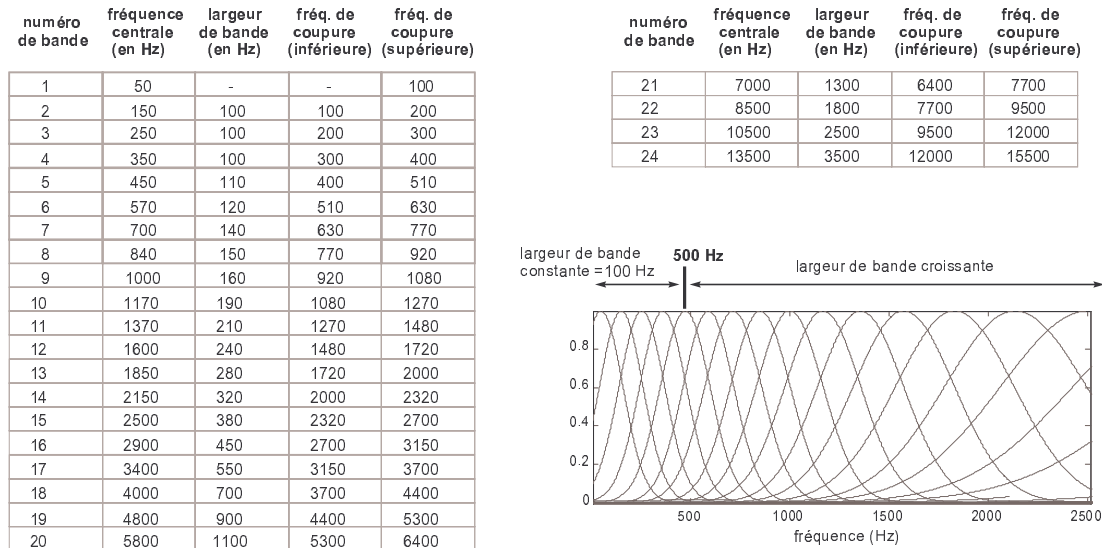


FIG. 6.13 – Liste des bandes critiques, déterminées par Zwicker et Feldtkeller. Le banc de filtres constitué des 15 premiers bandes et modélisé par une gaussienne.

Considérations Le découpage en échelle de Bark et plus particulièrement la position de la fréquence centrale sont, bien entendu, formels. En effet, comme le soulignent Zwicker et Feldtkeller, l'ouïe peut former une bande critique à n'importe quel point de l'échelle fréquentielle et réagencer le banc de filtres en fonction des bandes voisines. Cependant, l'échelle de Bark constitue une convention commode pour décrire le fonctionnement du système auditif. Il permet de constater que le système auditif possède une meilleure résolution fréquentielle dans les basses fréquences que dans les hautes fréquences. Une façon simple de construire le banc de filtres consiste à modéliser le profil des filtres à l'aide d'une gaussienne (**Fig. 6.13**) [96]. Toutefois, il faut garder à l'esprit que la gaussienne ne trace pas avec précision la forme des filtres auditifs, car les filtres auditifs ne sont pas symétriques : leurs flancs gauches sont plus raides que leurs flancs droits. D'une manière générale on peut observer quelques propriétés des filtres auditifs.

Unité de Bark L'unité de 1 Bark est définie de telle manière que chaque filtre auditif ait une largeur de 1 Bark. Par conséquent, on peut faire la cohérence entre l'échelle de fréquence et l'échelle de la résolution fréquentielle du système auditif. Ainsi, il est possible de convertir une fréquence en Hertz en sa valeur en Bark selon la formule

$$B(f) = 13 \cdot \arctan\left(\frac{0,76f}{1000}\right) + 3,5 \cdot \arctan\left(\left(\frac{f}{7500}\right)^2\right) \quad \text{Bark.} \quad (6.58)$$

Le fait que les bandes de Bark croissent d'une manière logarithmique les rapproche de la transformée en ondelettes. Afin de modéliser plus précisément la perception auditive, on peut utiliser une analyse de Gabor du signal avec une fenêtre constante au-dessous de 500 Hz et une analyse en ondelettes à partir de 500 Hz pour adapter la largeur de la fenêtre d'analyse d'une façon logarithmique.

Bark et compression audio Par ailleurs, le banc de filtres formé par les ondelettes, représente une assez bonne approximation de l'échelle perceptive. L'avantage des ondelettes est qu'elles bénéficient d'un algorithme de multirésolution qui est beaucoup plus efficace que l'implémentation directe des filtres correspondants à l'échelle de Bark. Par ailleurs, cette propriété des ondelettes est utilisée dans les algorithmes de compression audio. Les algorithmes se servent des spécificités psychoacoustiques de l'oreille et utilisent beaucoup de coefficients pour représenter fidèlement des bandes fréquentielles où l'oreille est sensible (surtout autour de 2-4 kHz) et un nombre réduit de coefficients pour les zones perceptives moins

¹en honneur du physicien allemand Barkhausen.

importantes. Pour le codage Mpeg sont utilisées 32 bandes critiques [97], alors que la compression Atrac de Minidisc (recentement arrêté de production) utilise 52 bandes en fonction des bandes critiques réparties en 20 sur les basses, 16 sur les moyennes et 16 sur les hautes fréquences [98].

6.5.2 ERB - Equivalent Rectangular Bandwidth

L'idée des bandes de Bark évolue vers une définition plus récente, celle des filtres ERB. Ce sont Moore et Glasberg qui introduisent une nouvelle mesure de la largeur des filtres auditifs en vue d'une estimation plus précise [99]. Ils définissent les ERB, bande rectangulaire équivalente (Equivalent Rectangular Bandwidth).

Expression Pour chaque filtre auditif, la valeur d'ERB est définie comme la largeur d'un filtre passe-bande avec la même fréquence centrale qui, en combinaison avec un bruit blanc, produirait la même énergie en sortie. Une expression analytique relie la largeur ERB (en Hz) et la fréquence centrale f_c

$$ERB_N(f_c) = 24,7 \left(\frac{4,37f_c}{1000} + 1 \right). \quad (6.59)$$

La largeur de chaque filtre augmente avec la fréquence et vaut environ 11-17 % de la valeur de la fréquence centrale.

ERB et fréquence en Hz D'une manière similaire à l'échelle de Barks, on peut définir, pour toute fréquence audible, une échelle des ERB, donc le numéro de bande s'obtient selon la formule

$$N_{ERB}(f_c) = q \log 10 \left(1 + \frac{f_c}{lq} \right), \quad (6.60)$$

où $l = 24,7$ et $q = 9.265$. L'expression de ERB peut être résumée comme suit en fonction de la fréquence centrale f_c

$$ERB = 6,23f_c^2 + 93,39f_c + 28,52 \quad (6.61)$$

La largeur de bande des ERB diffère de celle, définie dans les bandes critiques auparavant, surtout dans les basses fréquences, où la largeur des ERB est moindre que son équivalent dans la définition de Zwicker et Feldtkeller. Ceci est supposé donner une meilleure précision dans les basses fréquences.

6.5.3 Filtres gammatones

Les filtres de type gammatone sont actuellement les filtres auditifs les plus populaires pour modéliser l'action de la membrane basilaire. Les gammatones sont introduits par Holdsworth et Patterson au début des années 1990 [93] et sont implémentés numériquement par Slaney [100].

Fonction Gamma Les filtres gammatones correspondent aux réponses impulsionnelles des filtres cochléaires. Il faut mentionner que les filtres auditifs sont définis dans le domaine spectral, alors que les filtres gammatones sont exprimés dans le domaine temporel. Les filtres gammatones sont obtenus à l'aide de la modulation de la fonction gamma,

$$\Gamma(\gamma) = \int_0^{+\infty} e^{-t} t^{\gamma-1} dt \quad \text{où} \quad \Gamma(\gamma) = \Gamma(\gamma-1)! \quad , \gamma \in N, \quad (6.62)$$

et de la combinaison d'un polynôme croissant et d'une exponentielle décroissante.

Filtres gammatones Plus particulièrement, les filtres gammatones s'expriment dans le domaine temporel à l'aide de la formule

$$g(t) = At^{\gamma-1}e^{-2\pi Bt} \cos(2\pi f_0 t + \phi) \quad (6.63)$$

où γ représente l'ordre du filtre (typiquement d'ordre 4), f_0 est la fréquence centrale du filtre, ϕ est la phase et b est la largeur de bande fréquentielle, en ERB. Le spectre d'amplitude des filtres gammatone prend la forme

$$|G(f)| = \frac{A\Gamma(\gamma)}{(2\pi\sqrt{B^2 + (f - f_0)^2})^\gamma}. \quad (6.64)$$

En général, un petit nombre de filtres (environ 32) est suffisant pour couvrir la plage fréquentielle la plus importante, se situant entre 50 et 8000 Hz. Une illustration d'un banc de 10 filtres gammatones, couvrant la plage de [100-8000 Hz] est présentée sur la **Fig. 6.14**. Cet exemple utilise les paramètres par défaut de l'Auditory Toolbox de Slaney [100] et montre le profil asymétrique des filtres auditifs.

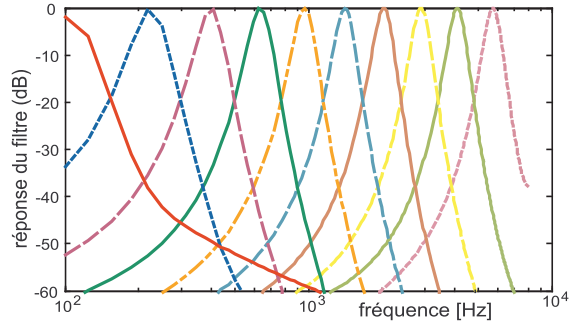


Fig. 6.14 – Banc de 10 filtres gammatones

6.5.4 Méthodes temps-fréquence intégrant des critères perceptifs

L'échelle de Bark, les ERB et les filtres gammatones sont trois évolutions des filtres auditifs qui possèdent un point commun : ils présentent une vision purement fréquentielle de la perception. Or, les phénomènes de masquage auditifs ont une manifestation aussi bien temporelle que fréquentielle. Cette caractéristique suggère une modélisation conjointe temps-fréquence du système auditif. Parmi les précurseurs d'une telle approche, on peut compter le modèle d'image auditive de Patterson et al. [101] et le modèle de l'excitation spectro-temporelle, étudié par Plack et Moore [102].

Distribution EarWig (EWD) L'approche temps-fréquence en échelle perceptive a commencé son développement dans les années 1990. Des critères perceptifs sont intégrés aux représentations linéaires comme la TFCT [103] ou les ondelettes [104]. Toutefois, les représentations linéaires ne permettent pas l'ajustement de la résolution à la fois en temps et en fréquence des filtres auditifs. La distribution EWD (Earwig distribution) est une application récente faite par O'Donovan qui intègre les filtres auditifs gammatones aux distributions bilinéaires de la classe de Cohen [105]. L'avantage de la classe de Cohen réside dans le noyau de lissage, dont la résolution peut être ajustée séparément en temps et en fréquence. Le modèle perceptif est imbriqué dans la construction du noyau de lissage de la EWD : le noyau de lissage en fréquence est obtenu à l'aide des filtres gammatone [106], alors que le lissage en temps est adapté sur la modélisation temporelle des filtres auditifs par Plack et Moore [102].

L'implémentation des filtres auditifs n'est pas une démarche très facile. La forme asymétrique des filtres auditifs et leur largeur qui augmente d'une manière non linéaire avec la fréquence, ne permettent pas l'utilisation des algorithmes rapides comme dans le calcul du spectrogramme ou le scalogramme. Le coût de calcul engendré est trop important et restreint sérieusement le champ d'utilisation des représentations temps-fréquence intégrant des critères perceptifs.

6.6 Applications pour les bruits moteur

Les bruits moteur forment une classe de signaux sonores qui se caractérise par la présence de structures à la fois impulsives et harmoniques. D'une manière générale, toutes les méthodes d'analyse temps-fréquence, qu'il s'agisse des représentations linéaires temps-fréquence (Gabor, TCFT) ou temps-échelle (ondelettes), des distributions bilinéaires de la classe de Cohen ou des méthodes intégrant des critères perceptifs, peuvent être utilisées dans l'analyse des bruits moteur. Cette section contient quelques exemples des méthodes décrites précédemment dans le contexte des bruits moteur. Je voudrais souligner que chacune de ces méthodes possède des avantages et des points faibles. Le choix d'une représentation dépendra des caractéristiques du signal analysé, de l'information recherchée, de la facilité d'interprétation, et enfin, d'une préférence personnelle envers l'utilisation d'une méthode plutôt que d'une autre.

6.6.1 Transformée de Gabor

La transformée de Gabor et la transformée de Fourier à court terme donnent une interprétation très intuitive du signal via le spectrogramme. De ce fait, le spectrogramme est un outil présent dans la quasi-totalité des applications d'analyse des signaux sonores.

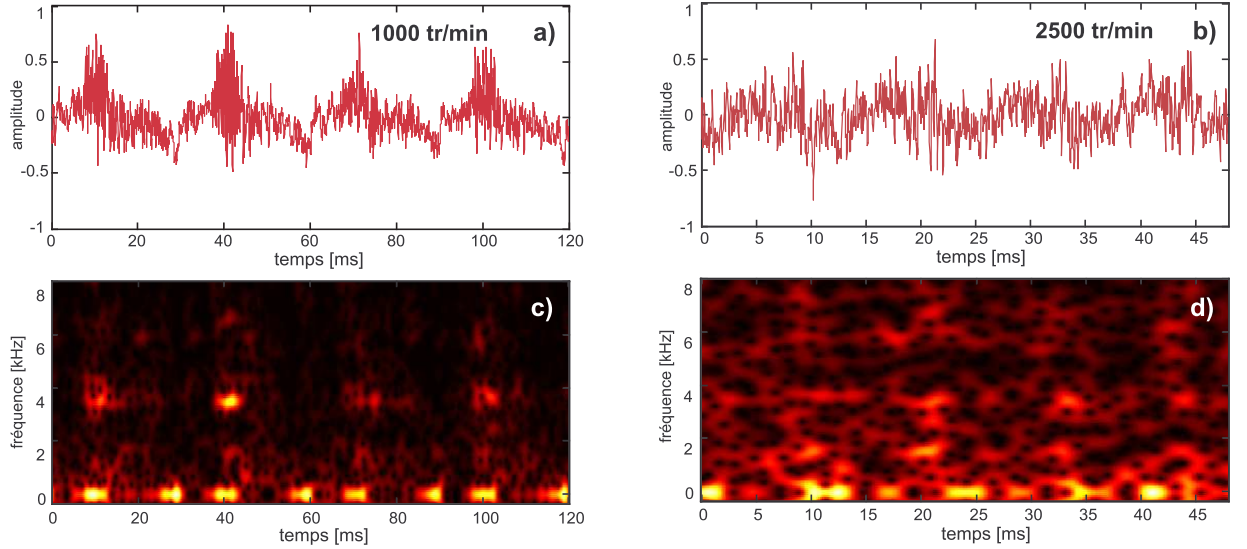


FIG. 6.15 – Les signaux temporels (a,b) et les spectrogrammes (c,d) de deux bruits diesel au banc. A gauche, à 1000 tr/min, à droite 2500 tr/min. Fenêtre d'analyse gaussienne ($\sigma=25$).

Afin de distinguer les différentes structures présentes dans le signal, on compare les spectrogrammes des bruits moteur à deux régimes différents, analysés avec la même fenêtre d'analyse, gaussienne d'une longueur de 256 échantillons et un paramètre $\sigma = 25$ (Fig. 6.15). Le bruit moteur à 1000 tr/min correspond à un ralenti élevé et possède une structure plutôt impulsive, bien localisée "verticalement" sur le spectrogramme, i.e en temps (Fig. 6.15, c). Le bruit moteur à un régime un peu plus élevé, à 2500 tr/min, laisse apparaître des caractéristiques plutôt harmoniques, bien localisée "horizontalement", en fréquence (Fig. 6.15, d).

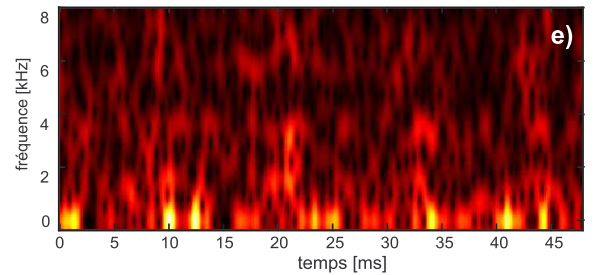


Fig. 6.16 – Spectrogramme du bruit à 2500 tr/min avec une fenêtre d'analyse gaussienne ($\sigma=10$).

Il faut mentionner que les propriétés de localisation du spectrogramme pour chacun de ces bruits dépendront des paramètres de la fenêtre d'analyse. Le même bruit moteur à 2500 tr/min, présenté sur la (Fig. 6.15, b) aura un spectrogramme localisé plutôt en fréquence avec une fenêtre dont le support en fréquence est petit ($\sigma = 25$, Fig. 6.15, d), et également un spectrogramme localisé plutôt en temps avec une fenêtre dont le support en temps est petit ($\sigma = 10$, Fig. 6.16, e).

6.6.2 Transformée en ondelettes

Dans le contexte automobile, la transformée en ondelettes continue est utilisée, en combinaison avec les méthodes d'autocovariance, pour obtenir les caractéristiques cycliques des données vibratoires [107]. L'analyse multirésolution de la transformée en ondelettes discrète est appliquée dans le traitement aux échelles différents des images captées par une caméra laser, filmant le processus d'injection diesel. [108].

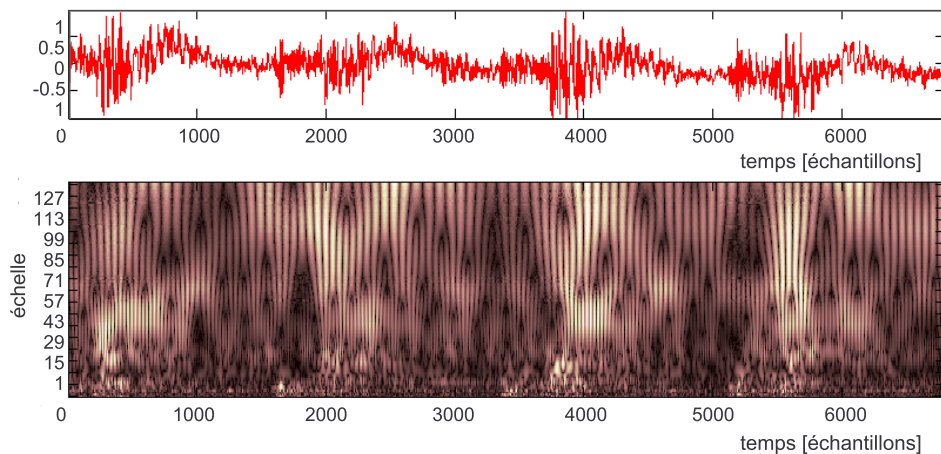


FIG. 6.17 – Le module de la transformée en ondelettes continue d’un cycle moteur au ralenti, ondelette d’analyse : Morlet

La Fig. 6.17 montre la transformée en ondelettes continue d’un cycle moteur au ralenti. L’analyse est faite avec Wavelet toolbox de Matlab, en utilisant l’ondelette de Morlet (de forme d’une gaussienne), sur une échelle allant de 1 à 128. Les petites échelles correspondent aux hautes fréquences, alors que les grandes échelles, aux basses fréquences. On peut observer l’élargissement du support fréquentiel avec l’augmentation de l’échelle. En résumé, la transformée en ondelettes peut fournir une image plus précise pour les sons impulsifs en hautes fréquences. Cependant, la notion temps-échelle semble moins naturelle que la notion temps-fréquence de la transformée de Gabor.

6.6.3 Echelle perceptive

Parmi les représentations, que je viens d’exposer, les représentations temps-fréquence en échelle perceptive sont les moins utilisées pour l’analyse des bruits moteur. Ceci peut s’expliquer par le fait, que ce sont des méthodes relativement récentes, actuellement en plein développement et pas encore bien établies. Les représentation en échelle perceptive visent plutôt les applications en physiologie de l’audition. De plus, l’estimation des filtres auditifs nécessite un temps de calcul supplémentaire très important. Par conséquent, seuls des signaux de très courte durée peuvent être analysés.

Exemple L’exemple d’analyse d’un quart de cycle moteur en échelle perceptive est donné sur la Fig. 6.18. La représentation est calculée en utilisant les paramètres par défaut du logiciel EWD (Ear-Wig Distribution), développé par O’Donovan [109]. Cette représentation tient compte de la perception de l’oreille. Toutefois, même sur un signal de courte durée (40 ms dans ce cas), le temps de calcul est trop important.

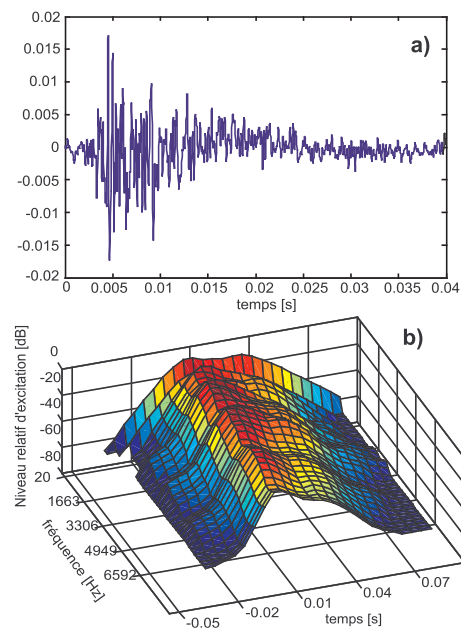


Fig. 6.18 – Un quart de cycle moteur : le signal temporel (a) et sa représentation EWD (b)

6.6.4 Distribution de Wigner-Ville

La distribution de Wigner-Ville est souvent utilisée dans sa forme lissée pour les applications sur les bruits moteur : pour la caractérisation fine en terme d'extraction de la signature sonore des différents parties moteur [110] ; dans l'étude de la cyclostationarité, les phénomènes stochastiques à l'intérieur du cycle moteur et pour le diagnostic de la combustion [7] ; dans l'estimation de la fréquence instantanée de résonance du cliquetis ("knock") sous différents conditions [111]. Une variante de la distribution de Wigner-Ville, la S-méthode [112], permet de représenter un signal multicomposantes à partir de la somme des distributions de pseudo-Wigner pour chaque signal séparément. Cette méthode est appliquée pour la détections des chocs dans les chambres de combustion.

Exemple La différence entre la distribution de Wigner-Ville et ses versions lissées est présentée sur la **Fig.6.19**. Elle analyse un quart de cycle moteur au ralenti. Le début du cycle ici est arbitraire et n'est pas synchronisé avec des données vilebrequin. L'analyse issue de la traditionnelle distribution de Wigner-Ville fait apparaître un événement dans le cycle moteur, sans doute la combustion, manifesté pendant les premiers 5 ms et autour de 4500 Hz. On observe une bonne concentration de l'énergie à l'endroit marqué par la combustion. En revanche, les termes d'interférences ne permettent pas de savoir quelles composantes sont présentes dans les autres régions du plan temps-fréquence. La distribution de pseudo-Wigner-Ville donne une image similaire, avec quelques différences dans la structure des interférences. C'est la version lissée de pseudo Wigner-Ville qui montre une information plus distincte des composantes. En revanche, la région la plus énergétique, associée à la combustion, manque de localisation et s'étale plus en temps et en fréquence que dans la représentation de Wigner-Ville non-lissée.

Discussion En pratique, l'analyse à l'aide de la distribution de Wigner-Ville impose l'utilisation de plusieurs images, lissées ou non. Les images non-lissées donnent avec précision l'instant et la fréquence d'apparition des événements. Les images, obtenues avec un lissage plus important permettent la visualisation des composantes d'un niveau plus faible, masquées auparavant par les interférences. Cet exemple laisse conclure que la version la plus "intéprétable", celle de pseudo Wigner-Ville lissée, se rapproche beaucoup en apparence du spectrogramme. Cependant, il faut garder à l'esprit l'avantage de la représentation de Wigner-Ville lissée qui reste la capacité d'ajuster indépendamment les résolutions en temps et en fréquence, ce qui n'est pas possible pour le spectrogramme. Par ailleurs, l'avantage de la distribution de Wigner-Ville était au départ la non-nécessité de choisir la fenêtre d'analyse. Or, le choix de la fenêtre s'impose lors d'un lissage.

Dans une telle interprétation, on considère judicieux d'utiliser la transformée de Gabor, plutôt que de chercher une version lissée de pseudo-Wigner-Ville qui va, finalement, se rapprocher des mêmes qualités de localisation.

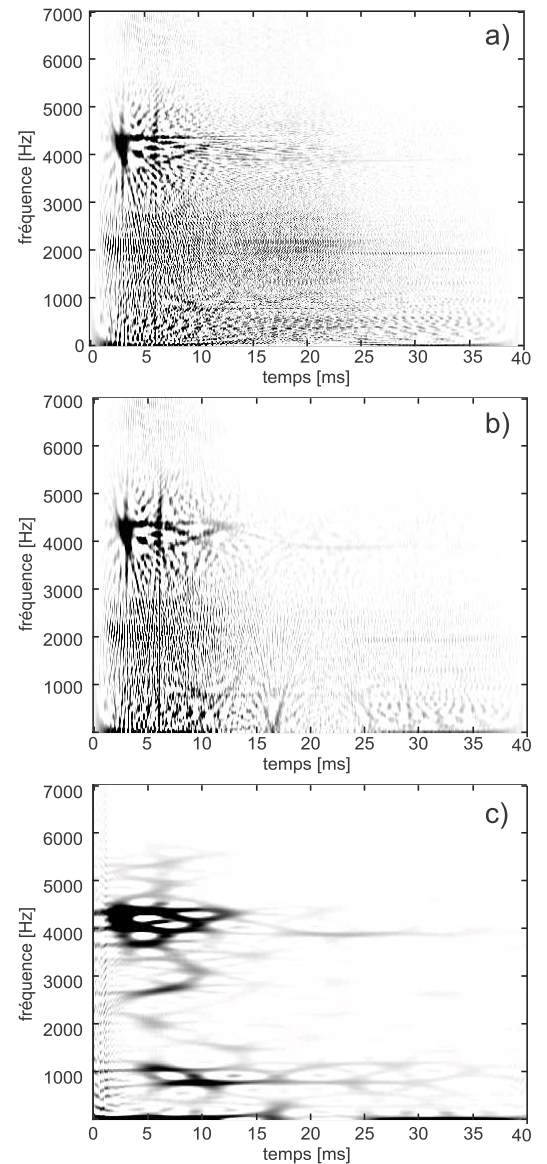


Fig. 6.19 – Un quart de cycle moteur diesel au ralenti : Wigner-Ville (a), pseudo Wigner-Ville (b) et pseudo Wigner-Ville lissée (c)

Conclusion sur les méthodes temps-fréquence

Dans ce chapitre j'ai exposé les principales méthodes d'analyse temps-fréquence du signal. Ils s'agit des représentations linéaires temps-fréquence (la Transformée de Gabor et la Transformée de Fourier à Court Terme) et temps-échelle (la Transformée en ondelettes), les distributions bilinéaires (Wigner-Ville, pseudo-Wigner-Ville) qui font partie d'une classe plus générique, la classe de Cohen et enfin, les représentations qui intègrent les particularités psychoacoustiques de l'oreille comme les filtres gammatones. Chacune de ces méthodes a été illustrée par des exemples qui montrent ses avantages et ses limitations. Une réflexion s'impose : il n'existe pas de méthode d'analyse universelle adaptée à tous les signaux. La méthode d'analyse doit être appropriée à la nature des signaux analysés. Il s'agit également d'un choix personnel, ainsi que des objectifs de la recherche. Toutes les méthodes d'analyse présentées ont été appliquées dans la littérature pour l'étude des bruits moteur.

Orientation de la thèse : J'ai préféré les représentations linéaires devant les distributions bilinéaires, car les représentations linéaires semblent plus appropriées en vue d'une synthèse granulaire du signal. J'ai également préféré la Transformée de Gabor devant la Transformée en ondelettes pour la simplicité et la clarté de travailler directement en termes de fréquences et non en termes d'échelle. Toutefois, la méthode d'analyse-synthèse, développée dans la thèse et exposée dans le dernier chapitre de cette partie peut être adaptée pour la transformée en ondelettes. Le choix de la Transformée de Gabor a rendu pour moi plus naturelle la comparaison avec la synthèse granulaire et la théorie des scènes auditives, ainsi que la construction des tests psychoacoustiques.

La réallocation - une concentration énergétique

La méthode de la réallocation a été créée pour répondre à la nécessité de concentrer l'énergie du signal. En fait, la précision du spectrogramme dans la localisation des composantes du signal dépend du support de la fenêtre d'analyse. Par ailleurs, d'après l'inégalité de Heisenberg-Gabor, cette localisation ne peut pas être infiniment petite. Plus le support sera petit en temps, plus il sera large en fréquence et vice-versa. La réallocation peut être vue alors comme un outil de post-traitement pour améliorer la lisibilité du spectrogramme et des représentations bilinéaires. Dans ce chapitre, je vais présenter les différentes formulations et moyens de calcul des opérateurs de la réallocation, en particulier, pour le spectrogramme. La comparaison entre la définition classique de la réallocation et des méthodes qui portent des idées similaires, comme la méthode d'"arête et squelette", permettra de situer la nouvelle méthode de réallocation, développée dans la thèse.

7.1 Opérateurs de la réallocation

La réallocation peut être considérée comme une méthode relativement récente. Elle est introduite par Kodera en 1976 afin d'améliorer la lisibilité du spectrogramme des signaux sismiques [90]. Pendant presque une vingtaine d'années, la réallocation n'a pas eu une utilisation répandue. Une relance de ce concept apparaît en 1995 dans les travaux de Auger et Flandrin, qui cherchaient un moyen de réduire les termes d'interférences des représentations bilinéaires [113]. Ils déduisent les opérateurs de la réallocation pour un large nombre de distributions de la classe de Cohen. Dans la continuité, une contribution majeure est apportée par Chassande-Mottin dans la définition de la réallocation différentielle [114].

7.1.1 Définition originale

Le point de départ de Kodera et col. est la "méthode de la fenêtre glissante", qui est une vision de la représentation temps-fréquence $S(t_0, f_0)$ du signal $x(t)$, analysé à travers d'un banc de filtres. La formulation, dans le domaine temporel ou fréquentiel est

$$S(t_0, f_0) = \int X(f)H(f - f_0)e^{i2\pi f t_0} df = e^{i2\pi f_0 t_0} \int x(t)h(t_0 - t)e^{i2\pi f_0 t} dt, \quad (7.1)$$

où $x(t)$ et $X(f)$ sont respectivement le signal analysé et son spectre de Fourier, alors que $h(t)$ et $H(f)$ sont respectivement la réponse impulsionnelle du filtre et sa fonction de transfert. Intuitivement, les variables t_0 et f_0 sont liées à la translation en temps et fréquence. Le module au carré $|S(t_0, f_0)|^2$ donne le spectrogramme du signal et Kodera fait remarquer que, dans l'expression équivalente en termes de densité d'énergie,

$$|S(t_0, f_0)|^2 = \iint E_x(t, f) \cdot E_h(t_0 - t, f - f_0) dt df, \quad (7.2)$$

la contribution énergétique du signal E_x est "lissée" par la réponse du filtre E_h . Ici sont utilisées les notations, mentionnant avec $*$ le complexe conjugué

$$E_x(t, f) = x(t)X^*(f)e^{i2\pi f t} \quad \text{et} \quad E_h(t, f) = h(t)H^*(f)e^{i2\pi f t} \quad (7.3)$$

La formule (7.2) montre qu'il n'est pas possible de séparer dans l'expression du spectrogramme l'énergie propre au signal de la contribution de la fenêtre d'analyse. C'est justement ce point faible du spectrogramme que la méthode de réallocation se propose d'améliorer.

Distribution énergétique et centres de gravité La proposition de Kodera consiste à attribuer la valeur du spectrogramme $|S(t_0, f_0)|^2$ non pas au centre du support de la fenêtre d'analyse, comme ceci est fait pour le spectrogramme, mais au point correspondant au centre de gravité des distributions énergétiques en temps et en fréquence (**Fig. 7.1**). Plus précisément, Kodera sépare d'abord dans l'intégrale

(7.2) la répartition énergétique en temps et en fréquence et définit deux distributions

$$\eta(f, t_0, f_0) = \int \operatorname{Re}[E_x(t, f)E_h(t_0 - t, f - f_0)]dt, \quad (7.4)$$

$$\xi(t, t_0, f_0) = \int \operatorname{Re}[E_x(t, f)E_h(t_0 - t, f - f_0)]df, \quad (7.5)$$

respectivement dans les domaines temporel et fréquentiel. Ici est pris en compte le fait que la partie imaginaire de (7.2) est nulle. Le support fréquentiel de la fenêtre d'analyse, localisée autour d'une fréquence f_0 , détermine un intervalle $[f_0 - \Delta f/2, f_0 + \Delta f/2]$.

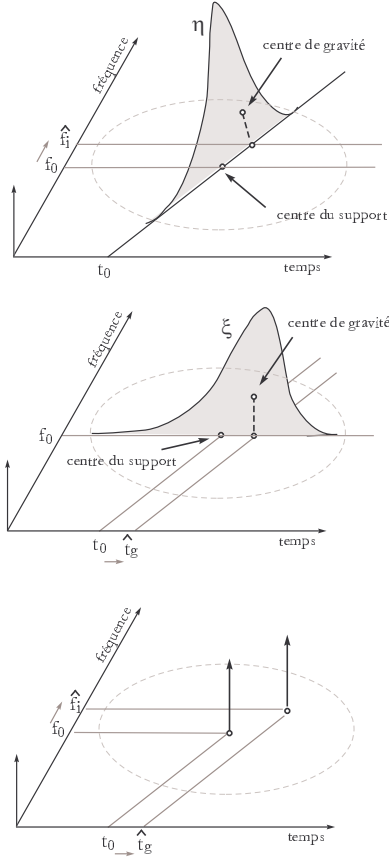


Fig. 7.1 – Réallocation aux centres de gravité des distributions en temps et en fréquence

Opérateurs de la réallocation La formulation des opérateurs de la réallocation à partir des centres de gravité, (7.7) et (7.6), exige un calcul très coûteux. Afin de formuler une expression plus convenable, Kodera utilise la forme complexe $S(t_0, f_0) = A(t_0, f_0)e^{i\Phi(t_0, f_0)}$ pour calculer les premiers moments (7.7) et (7.6). Ce calcul analytique introduit les dérivées partielles de la transformée $S(t_0, f_0)$, et aboutit à ce que je vais désigner comme "définition classique" des opérateurs de la réallocation

$$\hat{t}_g(t_0, f_0) = t_0 - \frac{1}{2\pi} \frac{\partial \Phi(t_0, f_0)}{\partial f_0}, \quad (7.8)$$

$$\hat{f}_i(t_0, f_0) = \frac{1}{2\pi} \frac{\partial \Phi(t_0, f_0)}{\partial t_0}. \quad (7.9)$$

Les opérateurs de la réallocation contiennent les expressions des dérivées de la phase $\frac{\partial \Phi}{\partial f_0}$ et $\frac{\partial \Phi}{\partial t_0}$. Ce fait témoigne que les valeurs réallouées \hat{t}_g et \hat{f}_i ont un lien avec les points stationnaires de la phase [72]. En effet, \hat{f}_i et \hat{t}_g sont reliés à la fréquence instantanée et le retard du groupe locaux.

Calcul par différences finies

La formulation classique des opérateurs passe par le calcul des

La fonction $\eta(f)$ possède un maximum dans cet intervalle, qui ne correspond pas forcément au point central f_0 . La coordonnée en fréquence de ce maximum va mieux représenter l'énergie de l'intégrale et l'énergie de l'intégrale doit alors être attribuée à cette fréquence. Similairement est déterminée la meilleure position $\xi(t)$ en temps. L'action de déplacer l'énergie vers la "vraie" valeur temporelle ou fréquentielle exprime le processus de la réallocation. Un moyen approprié de trouver les "vraies" coordonnées en temps et fréquence, est de calculer les premiers moments normalisés des distributions $\eta(f)$ et $\xi(t)$, respectivement

$$\hat{f}_i(t_0, f_0) = \frac{\int f \cdot \eta(f, t_0, f_0) df}{\int \eta(f, t_0, f_0) df} \quad (7.6)$$

et

$$\hat{t}_g(t_0, f_0) = \frac{\int t \cdot \xi(t, t_0, f_0) dt}{\int \xi(t, t_0, f_0) dt} \quad (7.7)$$

Les équations (7.7) et (7.6) définissent les opérateurs de réallocation. Les points $\hat{f}_i(t_0, f_0)$ et $\hat{t}_g(t_0, f_0)$ portent la signification physique des centres de gravité des distributions $\eta(f)$ et $\xi(t)$. La valeur de $|S(t_0, f_0)|^2$ doit être déplacée du point (t_0, f_0) vers le nouveau point (\hat{t}_g, \hat{f}_i) . Cette méthode est appelée par Kodera "modified moving window method" (méthode modifiée de la fenêtre glissante), en association de la représentation temps-fréquence qui est, quant à elle, la "méthode de la fenêtre glissante".

dérivées de la phase de la transformée. Dans l'article de Kodera, les opérateurs de la réallocation sont calculés numériquement à l'aide des différences finies,

$$\hat{t}_g(t_0, f_0) = t_0 - \frac{\Phi(t_0, f_0 + \Delta f) - \Phi(t_0, f_0)}{2\pi\Delta t}, \quad (7.10)$$

$$\hat{f}_i(t_0, f_0) = \frac{\Phi(t_0 + \Delta t, f_0) - \Phi(t_0, f_0)}{2\pi\Delta t}, \quad (7.11)$$

où dans l'algorithme original est d'abord calculée la coordonnée \hat{f}_i et ensuite \hat{t}_g . La valeur du point recherché (\hat{t}_g, \hat{f}_i) est trouvé en une seule itération. L'utilisation des différences finies semble très pratique d'un point de vue numérique. Cependant, dans certains cas, l'erreur d'approximation peut devenir importante. De plus, les discontinuités de la phase peuvent provoquer une instabilité numérique.

La méthode de réallocation de Kodera assure une localisation des composante du signal beaucoup plus précise que le spectrogramme traditionnel. Elle est comparée avec plusieurs méthodes afin de montrer ses meilleurs résultats en termes de concentration d'énergie des composantes du signal [115]. Cependant, cette méthode ne bénéficie pas d'un succès immédiat, probablement à cause du manque d'une preuve théorique de son efficacité, et de certains inconvénients de l'implémentation qui peuvent être engendrés lors du calcul des dérivées de la phase.

7.1.2 Généralisations

L'idée de Kodera est reprise beaucoup plus tard par Auger et Flandrin dans le but de trouver un moyen de réduire les termes d'interférences qui sont générées dans les distributions bilinéaires. Comme définie précédemment, la généralisation de ces représentations du signal est donnée par la classe de Cohen

$$C(x, t, \omega) = \int \int \phi_{TF}(u, \Omega) WV(x, t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}, \quad (7.12)$$

où les éléments de cette classe sont obtenus par lissage de la distribution de Wigner-Ville

$$WV(x; t, \omega) = \int x(t + \tau/2) x^*(t - \tau/2) e^{-i\omega\tau} d\tau, \quad (7.13)$$

avec un noyau de lissage ϕ_{TF} . Les exemples, donnés dans le chapitre précédent ont montré que, effectivement, le lissage réduit les termes d'interférence, mais en même temps rend floues les composantes du signal.

Opérateurs de réallocation, classe de Cohen La méthode, appelée par Auger et Flandrin comme méthode de "réallocation", est tout à fait dans le même esprit que le procédé introduit précédemment par Kodera. Elle consiste à attribuer l'énergie non au centre du noyau de lissage, mais au centre de gravité de ces contributions énergétiques. Les formules de réallocation pour la classe de Cohen, prennent la forme suivante [116]

$$\hat{t}(x; t, \omega) = t - \frac{\int \int u \cdot \phi_{TF}(u, \Omega) WV(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}}{\int \int \phi_{TF}(u, \Omega) WV(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}}, \quad (7.14)$$

$$\hat{\omega}(x; t, \omega) = \omega - \frac{\int \int \Omega \cdot \phi_{TF}(u, \Omega) WV(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}}{\int \int \phi_{TF}(u, \Omega) WV(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}}. \quad (7.15)$$

En s'appuyant sur les formules généralisées (7.14) et (7.15), Auger et Flandrin déduisent les opérateurs de la réallocation pour quelques distributions couramment utilisées comme Pseudo-Wigner-Ville, Margenau-Hill, ainsi que des représentation temps-échelle comme le scalogramme. Auger et Flandrin remarquent que la réallocation fait perdre les propriétés bilinéaires des distributions de la classe de Cohen, cependant, elle s'avère plus appropriée pour atténuer les termes d'interférences que le simple lissage. Outre ces caractéristiques, le grand avantage de définir les opérateurs de la réallocation dans le contexte de la classe de Cohen, se concrétise dans une expression très pratique pour le calcul des opérateurs de la réallocation du spectrogramme.

Réallocation du spectrogramme Le spectrogramme est défini dans la classe de Cohen, comme une distribution, dont le noyau de lissage est la distribution de Wigner-Ville d'une fenêtre d'analyse $h(t)$. Le spectrogramme est défini dans la classe de Cohen comme

$$S_h(x; t, \omega) = \int \int WV(h; u, \Omega) WV(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}. \quad (7.16)$$

Effectivement, la distribution $S_h(x; t, \omega)$ est équivalente au module au carré de la Transformée de Fourier à Court Terme (TFCT), et on peut le noter

$$S_h(x; t, \omega) = |TCFT_h(x; t, \omega)|^2 \quad \text{où} \quad TCFT_h(x; t, \omega) = \int x(t) \cdot h^*(t - u) e^{-i\omega u} du. \quad (7.17)$$

Le grand avantage de l'interprétation du spectrogramme dans la classe de Cohen se manifeste à travers l'expression des opérateurs de la réallocation, déduite par Auger et Flandrin. Cette expression permet le calcul des opérateurs directement à partir du module au carré de la TFCT [113]

$$\hat{t}(x; t, \omega) = t - \Re \left\{ \frac{TFCT_{th}(x; t, \omega) \cdot TFCT_h^*(x; t, \omega)}{|TFCT_h(x; t, \omega)|^2} \right\}, \quad (7.18)$$

$$\hat{\omega}(x; t, \omega) = \omega - \Im \left\{ \frac{TFCT_{dh}(x; t, \omega) \cdot TFCT_h^*(x; t, \omega)}{|TFCT_h(x; t, \omega)|^2} \right\}, \quad (7.19)$$

où $TFCT_h$, $TFCT_{dh}$ et $TFCT_{th}$ sont les spectrogrammes calculés avec trois fenêtres d'analyse différentes : la fenêtre d'analyse $h(t)$, sa dérivée $dh(t)/dt$ et la fenêtre $t.h(t)$. Un exemple des trois fenêtres obtenues à partir de la gaussienne est montré sur la **Fig. (7.2)**. Par ailleurs, dans le cas de la gaussienne, les fenêtres $dh(t)/dt$ et $t.h(t)$ sont proportionnelles. Le bénéfice des formules (7.18) et (7.19) est que le calcul des opérateurs de la réallocation ne passe plus par la phase de la transformée. Par ailleurs, ces formules peuvent être utilisées également pour la transformée de Gabor.

Formules simplifiées Il est possible de simplifier les formules (10.32) et (10.33), en fournissant une expression encore plus compacte des opérateurs de réallocation du spectrogramme

$$\hat{t}(x; t, \omega) = t - \Re \left\{ \frac{TFCT_{th}(x; t, \omega)}{TFCT_h(x; t, \omega)} \right\} \quad (7.20)$$

$$\hat{\omega}(x; t, \omega) = \omega - \Im \left\{ \frac{TFCT_{dh}(x; t, \omega)}{TFCT_h(x; t, \omega)} \right\} \quad (7.21)$$

Les formules (7.20) et (7.21) assurent l'expression exacte des opérateurs de réallocation définis par Kodera. Désormais, le procédé est dispensé de l'utilisation des différences finies pour calculer les dérivées de la phase. Par ailleurs, l'utilisation des différences finies est accompagnée toujours d'une erreur d'approximation. Ici cette erreur est évitée.

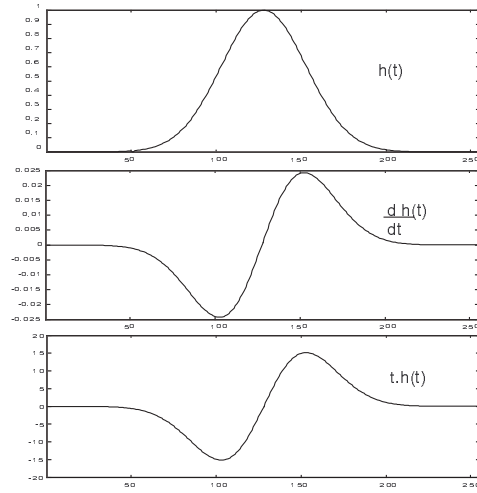


Fig. 7.2 – Les trois fenêtres, $h(t)$, $\frac{dh(t)}{dt}$ et $t.h(t)$ obtenues à partir d'une gaussienne

Avantages et limites La réallocation du spectrogramme donne une localisation parfaite pour les impulsions et les chirps linéaires. Un inconvénient majeur de la réallocation est qu'elle ne donne pas la possibilité d'une reconstruction parfaite du signal. En effet, en déplaçant l'énergie, l'information contenue dorénavant sur la phase du signal est perdue, ce qui empêche de revenir au signal original. Ce point faible de la réallocation a été certainement la raison pour laquelle elle a été ignorée pour des applications comme le codage des signaux audio où la propriété de reconstruction est davantage exigée.

7.1.3 Réallocation différentielle

Dans une étude très complète sur la réallocation, Chassande-Mottin considère les opérateurs de la réallocation comme un champ de vecteurs qui se déplace dans le plan temps-fréquence. A l'aide de cette

formulation vectorielle, il est possible de comprendre les principes selon lesquels s'effectue le déplacement de l'énergie lors de la réallocation. Par ailleurs, la géométrie du champ de réallocation est déterminée aussi par l'existence de points stables, qui se comportent comme des "attracteurs" d'énergie.

Champs de vecteurs Le champ généralise des vecteurs de la réallocation $\hat{r}(t, \omega)$ s'écrit sous une forme complexe les deux opérateurs de réallocation en temps (partie réelle) et en fréquence (partie imaginaire) et peut être noté également dans la forme $r_x(t, \omega)$

$$\hat{r}(t, \omega) = (\hat{t}_x - t) + i(\hat{\omega}_x - \omega) \quad \text{ou} \quad r_x(t, \omega) = \left(\frac{\hat{t}_x - t}{\Delta t}, \frac{\omega_x - \omega}{\Delta \omega} \right) \quad (7.22)$$

Ici $r_x(t, \omega)$ est normalisé par les largeurs temporelle Δt et fréquentielle $\Delta \omega$ du noyau ϕ_{TF}

$$\Delta t = \left(\int t^2 \phi_{TF}(t, \omega) \frac{dtd\omega}{2\pi} \right)^{1/2} \quad \text{et} \quad \Delta \omega = \left(\int \omega^2 \phi_{TF}(t, \omega) \frac{dtd\omega}{2\pi} \right)^{1/2} \quad (7.23)$$

Dans la continuation de cette définition, les opérateurs de la réallocation pour le spectrogramme déduits par Auger et Flandrin, 7.20 et 7.21, s'écrivent à l'aide du vecteur de déplacement normalisé sous la forme

$$r(t, \omega) = \frac{1}{\Delta t_h} \text{Re} \left\{ \frac{TCFT_{th}}{TCFT_h} \right\} - \frac{i}{\Delta \omega_h} \text{Im} \left\{ \frac{TCFT_{dh}}{TCFT_h} \right\} \quad (7.24)$$

où Δt_h et $\Delta \omega_h$ décrivent le support temporel et fréquentiel des la fenêtre h .

Fenêtre gaussienne Dans le cas d'une fenêtre gaussienne, que Chassande-Mottin définit sous la forme $h(t) = \pi^{-1/4} \sigma^{-1/2} e^{-t^2/2\sigma^2}$, les fenêtres $t.h(t)$ et dh/dt sont proportionnelles. Par conséquent, l'expression du champ des vecteurs (7.24) prend la forme très simple

$$\frac{dh(t)}{dt} = -\frac{1}{\sigma^2} t.h(t) \quad \implies \quad r_x^h = \frac{\sqrt{2}}{\sigma} \frac{TCFT_{th}}{TCFT_h} \quad (7.25)$$

La réallocation différentielle a été liée par Chassande-Mottin à la mesure d'information du signal et le partitionnement du plan temps-fréquence en fonction des composantes du signal.

7.2 Interprétation de la réallocation

Dans la définition classique, donnée par Kodera, les opérateurs de la réallocation s'expriment en termes des dérivées de la phase de la transformée de Gabor, liées elles mêmes à la notion de fréquence instantanée et retard du groupe.

7.2.1 Lien avec la fréquence instantanée

Les nouvelles coordonnées en temps et fréquence \hat{t}_g et $\hat{\omega}_i$, déterminées par les opérateurs de la réallocation (7.7) et (7.6), traduisent une idée similaire de la fréquence instantanée et du retard de groupe d'un signal. Dans ces deux cas, il s'agit des grandeurs, définies au sens d'un barycentre. En effet, la fréquence instantanée d'un signal, obtenue comme la dérivée en temps de la phase du signal Φ_s , porte également le sens physique d'un barycentre fréquentiel. De même, dans le domaine fréquentiel, le retard du groupe exprime le barycentre en temps du signal

$$\omega_i(t) = \frac{d\Phi_s(t)}{dt} = \text{Re} \frac{\int \omega \cdot \hat{s}(\omega) e^{i\omega t} d\omega}{\int \hat{s}(\omega) e^{i\omega t} d\omega}, \quad (7.26)$$

$$\bar{t}_g(\omega) = -\frac{d\Phi_s(\omega)}{d\omega} = \text{Re} \frac{\int t \cdot s(t) e^{-i\omega t} dt}{\int s(t) e^{-i\omega t} dt}, \quad (7.27)$$

L'opérateur de réallocation en temps peut être interprété comme le retard de groupe, vu à travers la fenêtre d'analyse $h(t)$. De même, l'opérateur de réallocation en fréquence peut être vu comme la fréquence instantanée du signal filtré par une bande de fréquence, définie par la fenêtre d'analyse.

7.2.2 Lien avec les méthodes "arête et squelette"

L'interprétation des opérateurs de la réallocation en termes de fréquence instantanée et de retard de groupe rend visible le lien de la réallocation avec la méthode de type "arête et squelette". Cette méthode a été développée par Guillemain et Kronland-Martinet pour les signaux musicaux [71].

Principe La méthode d'arête et squelette consiste à détecter certaines lignes importantes dans le plan temps-fréquence, appelées des arêtes. Les arêtes sont relatives à la loi d'évolution de la fréquence instantanée ou du retard du groupe. La restriction de la transformée de Gabor ou la transformée en ondelettes le long des arêtes donne le squelette. Le squelette est une représentation simplifiée et non redondante qui a l'avantage de permettre la reconstruction du signal. En comparaison, les opérateurs de la réallocation ne sont pas destinés à la reconstruction du signal.

Algorithme de point fixe L'estimation de l'arête est en lien direct avec l'estimation des lois de la fréquence instantanée et du retard du groupe du signal. L'arête peut être extraite à partir de la phase de la transformée à l'aide d'un algorithme de point fixe. L'arête est estimée en termes de fréquence instantanée de manière que la fréquence α , à un paramètre τ_0 fixe, satisfasse la relation

$$\frac{\partial \Phi(\tau_0, \alpha)}{\partial \tau} = \alpha. \quad (7.28)$$

L'algorithme consiste à se poser dans une valeur initiale en fréquence α_0 et ensuite d'itérer l'équation 7.28 de façon que, à chaque nouvelle itération i , une nouvelle fréquence α_{i+1} soit estimée afin de satisfaire la relation

$$\alpha_{i+1}(\tau_0) = \frac{\partial \Phi(\tau_0, \alpha_i(\tau_0))}{\partial \tau}. \quad (7.29)$$

L'algorithme de point fixe s'arrête quand la différence entre deux valeurs de α_i et α_{i+1} devienne trop petite. Les points de l'arête correspondent aux points où l'opérateur de réallocation en fréquence est stationnaire. Ces points déterminent l'arête "horizontale". L'estimation de l'arête en termes du retard de groupe détermine, à un paramètre α_0 fixe, un ensemble de points qui satisfont la relation

$$\frac{\partial \Phi(\tau, \alpha_0)}{\partial \alpha} = 0. \quad (7.30)$$

D'une manière similaire à l'itération de l'équation 7.28, l'adaptation de l'algorithme de point fixe pour l'équation 7.30 donne l'arête "verticale" dont les valeurs sont estimées après plusieurs itérations de

$$\tau_{i+1}(\alpha_0) = \tau_i(\alpha_0) - \frac{\partial \Phi(\tau_i(\alpha_0), \alpha_0)}{\partial \alpha}, \quad (7.31)$$

jusqu'à ce que la différence entre τ_i et τ_{i+1} devienne négligeable.

Démarches qualitative et quantitative La comparaison entre les opérateurs de la réallocation et la méthode d'arête et squelette donne lieu à une interprétation très importante, qui permettra de situer la nouvelle méthode, développée dans la thèse. Cette interprétation distingue deux approches complémentaires dans la concentration de l'énergie du signal : qualitative et quantitative.

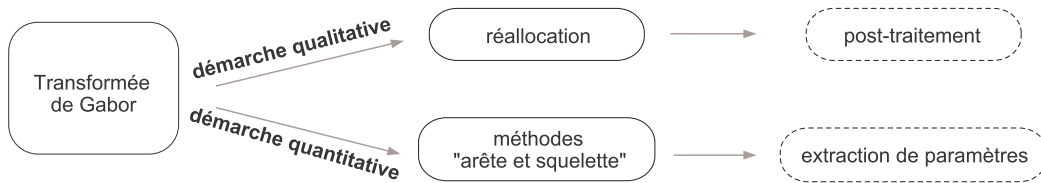


FIG. 7.3 – Démarche qualitative et quantitative de la réallocation et "arête et squelette"

La démarche qualitative consiste à rendre plus nette l'image temps-fréquence, en déplaçant l'énergie du signal. Cette opération peut être vue comme un post-traitement du spectrogramme et ne permet pas une reconstruction du signal. Au contraire, l'approche quantitative a pour objectif l'estimation des paramètres du signal en vue d'une resynthèse (**Fig. 7.3**). Ces deux démarches, bien que différentes dans leurs objectifs, partagent les mêmes outils, et plus particulièrement, l'utilisation des dérivées partielles de la phase de la transformée.

7.3 Applications et exemples

La réallocation permet une concentration plus précise de l'énergie du signal et rend l'image temps-fréquence plus nette. La réallocation est utilisée principalement pour le spectrogramme, mais également pour améliorer la lisibilité des versions lissées des distributions bilinéaires de type Wigner-Ville.

Spectrogramme Dans le cas des représentations linéaires, dont la TCFT est l'exemple classique, la réallocation réduit l'influence du support de la fenêtre d'analyse (**Fig. 7.4**). La concentration de l'énergie du signal se fait le long d'une ligne dans le plan temps-fréquence, comparable aux lignes de la fréquence instantanée ou du retard de groupe.

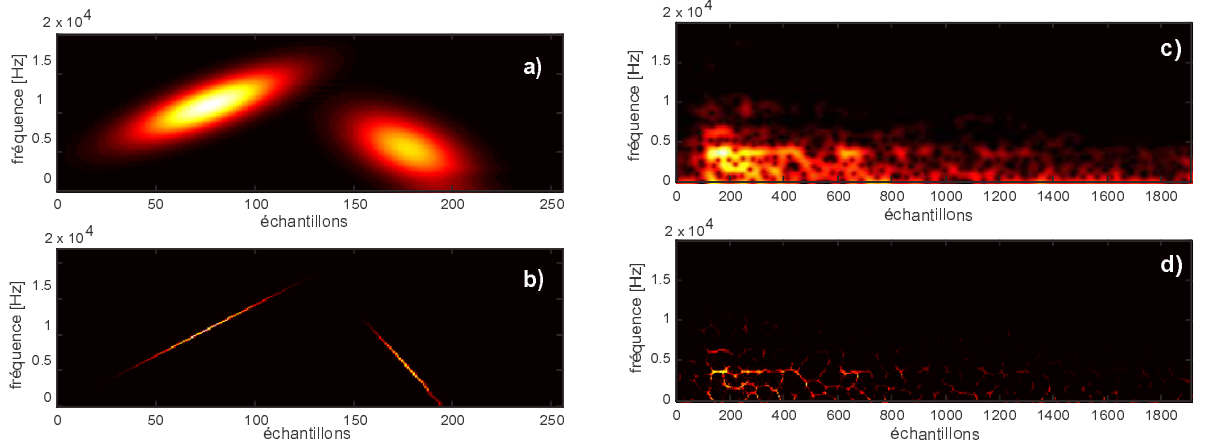


FIG. 7.4 – Les spectrogrammes (en haut) et les spectrogrammes réalloués selon Kodera (en bas) de deux atomes gaussiens (à gauche) et un quart de cycle moteur au ralenti (à droite)

Pour un chirp linéaire à enveloppe gaussienne, la réallocation déplace l'énergie du signal le long d'une droite dans le plan temps-fréquence penchée à un angle qui dépend du paramètre du chirp (ou encore le paramètre d'orientation dans le plan temps-fréquence) (**Fig. 7.4 (a,b)**). On proposera par la suite de nouveaux opérateurs de réallocation qui permettront de concentrer l'énergie du chirp non plus sur une droite, mais sur un point, en gardant l'information des paramètres du support et d'orientation du chirp. La réallocation d'un signal plus complexe (un quart de cycle moteur au ralenti, **Fig. 7.4 (c,d)**) montre aussi que les opérateurs de réallocation réduisent l'influence de la fenêtre d'analyse.

Pseudo Wigner-Ville La réallocation permet d'atténuer également l'effet de la fenêtre de lissage dans le cas des représentations bilinéaires comme le pseudo Wigner-Ville. En fait, comme déjà discuté, les versions lissées réduisent les termes d'interférences de la distribution de Wigner-Ville, mais en même temps "étalent" l'information du signal. La réallocation de l'image temps-fréquence de la pseudo Wigner-Ville permet, d'une manière similaire à la réallocation du spectrogramme, de diminuer l'influence du support de la fenêtre de lissage.

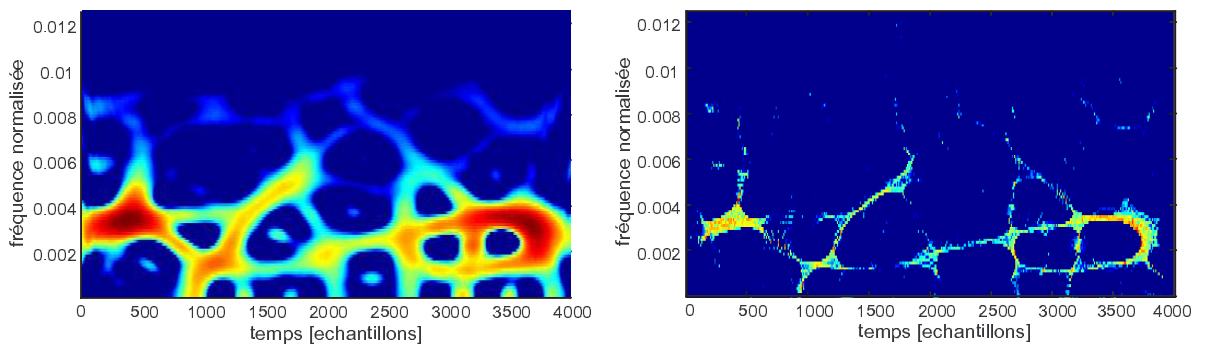


FIG. 7.5 – Bruit de changement de rapport de la boîte de vitesse : la représentation de pseudo Wigner-Ville (à gauche) et sa version réallouée (à droite)

Conclusion sur la réallocation

La réallocation est une méthode relativement récente, qui peut être vue comme un post-traitement des représentations temps-fréquence du signal. Son utilité est évidente, puisqu'elle permet une image plus nette du spectrogramme et des versions lissées des distributions bilinéaires. Les recherches dans les derniers vingt ans ont proposé, par rapport à la définition initiale, des moyens plus pratiques de calcul des opérateurs de la réallocation. On va utiliser par la suite une manière rapide de calculer les opérateurs pour le spectrogramme, qui réside dans l'utilisation de la TFCT avec trois fenêtres différentes. Cette manière de calculer les opérateurs est plus difficilement applicable pour la transformée en ondelettes. Une démarche qui généralise les opérateurs de la réallocation est donnée par la réallocation différentielle. La reconstruction du signal à partir de la réallocation, n'est pas possible. Toutefois, des méthodes s'inspirant des mêmes idées que la réallocation, comme la méthode l'"arête et squelette", ont été développées dans l'objectif de l'estimation des paramètres du signal en vue de sa reconstruction.

Orientation de la thèse : La méthode de la réallocation est une base importante de l'algorithme d'analyse-synthèse. Notre démarche combinera l'idée qualitative de la réallocation et l'idée quantitative de la méthode d'"arête et squelette". La concentration de l'énergie du signal sera encore plus fine, au sens où elle sera non plus sur des lignes de la fréquence instantanée (ou le retard du groupe), comme le font les opérateurs de la réallocation, mais sera localisée sur des points temps-fréquence. L'approche quantitative de la nouvelle méthode consistera à estimer les paramètres du signal, définis par le modèle granulaire. De plus, la méthode d'analyse-synthèse va intégrer l'algorithme itératif de point fixe, prolongeant les méthodes d'"arête et squelette", afin d'ajuster les positions temps-fréquence des atomes. Par ailleurs, le fait de concentrer l'énergie du signal sur un nombre restreint d'atomes temps-fréquence, soulève la problématique des décompositions atomiques et parcimonieuses du signal.

Algorithmes de décomposition

Les méthodes de décomposition occupent une place importante dans le domaine du traitement des signaux. La décomposition fournit une approximation du signal sous la forme d'une somme de fonctions issues de la même famille. La décomposition est une approche mathématique qui apporte une représentation compacte et souvent implicite du signal. Les séries de Fourier, les atomes temps-fréquence de Gabor et les ondelettes sont des exemples communs de familles de décomposition. La problématique des décompositions du signal comporte deux questions essentielles : comment approximer au mieux le signal et comment l'approximer au mieux avec peu d'éléments ? Ce chapitre rappelle quelques algorithmes développés lors des dernières années, comme les méthodes de type "poursuit", ou d'autres méthodes s'appuyant sur une modélisation probabiliste des coefficients.

8.1 Problématique

Le principal objectif des algorithmes de décomposition est d'approximer le signal $s(t)$ par une combinaison de fonctions élémentaires $g_k(t)$, tel que

$$s(t) \sim \sum_{k=1}^M a_k \cdot g_k(t). \quad (8.1)$$

Les fonctions sont choisies parmi un grand nombre de fonctions de la même famille. Les sinusoides, les fonctions de Gabor (gaborettes), les ondelettes, les chirps linéaires sont des exemples de fonctions élémentaires.

Classification des méthodes Il existe deux grandes classes de méthodes d'approximation des signaux. La première classe comprend l'approximation linéaire du signal sur une base orthogonale pré-définie. Un tel exemple est la base de Karhunen-Loève qui est obtenue à l'aide des vecteurs propres de l'opérateur de corrélation [117]. En effet, il s'agit d'une analyse en composantes principales. Cette approche a l'avantage de fournir une décomposition unique du signal. Cependant, la base orthogonale semble trop rigide par rapport à la nature des signaux divers, qui peuvent comporter à la fois des structures tonales et impulsives. La deuxième classe de méthodes introduit les approximations non-linéaires qui utilisent une base adaptative, souvent très redondante [117]. Cette base est déterminée au fur et à mesure de l'analyse du signal, elle est plus souple par rapport à la structure locale du signal et donne de meilleures approximations. En revanche, comme la solution n'est plus unique, la difficulté principale surgit dans la construction de la décomposition optimale. Cette solution est proposée par diverses méthodes de type poursuit (matching, basis, tree ..) ou la théorie des frames.

Approximations non-linéaires Les approximations non-linéaires avec des fonctions élémentaires font l'objet de trois caractéristiques importantes : la justification du modèle, le choix des fonctions élémentaires et le critère de sélection des éléments qui participent dans l'approximation finale du signal. Ces caractéristiques sont déterminées en fonction de la nature des signaux analysés, de la précision d'approximation voulue, du coût de calcul. En règle générale, on veut garder un minimum de coefficients, tout en conservant l'énergie du signal. A partir de ces considérations, plusieurs algorithmes sont développés, à des fins diverses en traitement de signal (détection, débruitage, compression [118], modélisation, apprentissage automatique, séparation de sources), et sont appliqués pour la parole, les signaux musicaux, les bruits impulsifs.

Parcimonie On a déjà remarqué, que les fonctions élémentaires de Gabor introduisent un grand degré de redondance, surtout quand on ne veut pas sacrifier leurs bonnes propriétés de localisation. Le fait de choisir juste un nombre limité de fonctions, largement inférieur au nombre total de membres de la famille de fonctions, nous amène à une approximation dite parcimonieuse (en anglais, "sparse decomposition"), i.e qui contient peu d'éléments.

8.2 Approximations adaptatives

Les approximations parcimonieuses ont pour but de décomposer le signal seulement sur un nombre restreint de coefficients, tout en gardant un maximum d'information qu'il contient. Ceci revient à chercher un sous-dictionnaire du dictionnaire redondant. Un dictionnaire représente un ensemble de fonctions $D = (g_\gamma : \gamma \in \Gamma)$ dont la combinaison linéaire est dense dans l'espace du signal H . Le plus petit dictionnaire possible forme une base de H . Le paramètre γ est un indice de la fonction élémentaire qui peut avoir un sens de paramètre de localisation temps-fréquence, de paramètre d'échelle, etc. Dans un dictionnaire redondant, la représentation du signal n'est plus unique.

8.2.1 Optimalité

Les décompositions atomiques cherchent à approximer un signal avec M -termes (atomes) issus d'un dictionnaire.

$$s(t) = \sum_{k=1}^M a_k \cdot g_k(t) + \varepsilon. \quad (8.2)$$

L'objectif de la décomposition est de minimiser l'erreur d'approximation ε , souvent mesurée au sens quadratique.

Optimalité et redondance La question principale des approximations adaptatives est de choisir les fonctions élémentaires afin d'obtenir une décomposition optimale. Par ailleurs, il est montré par Davis, que la recherche de l'approximation optimale dans un dictionnaire redondant est un problème NP-complet [119]. Par conséquent, d'un point de vue mathématique, il est impossible de trouver l'approximation optimale, mais seulement une solution sous-optimale.

Optimalité et perception En réalité, en prenant en compte des spécificités de l'oreille humaine, il est judicieux de poser la question suivante : est-ce que l'approximation optimale avec une minimisation de l'erreur au sens quadratique fournit la décomposition la plus satisfaisante en termes perceptifs pour l'oreille ? C'est un champ scientifique très vaste. Cependant, en partant de l'hypothèse que peut-être l'oreille ne prend pas en compte une partie de l'information du signal, on peut espérer obtenir des décompositions sous-optimales qui se rapprochent suffisamment de l'"optimalité" en termes perceptifs. D'ailleurs, dans les algorithmes de codage audio, certaines propriétés de l'oreille, comme par exemple la sensibilité dans la plage fréquentielle [3-8 KHz], sont utilisées pour approximer finement le signal dans cette zone pour satisfaire les exigences perceptives, tout en réservant moins d'informations dans les parties jugées moins importantes pour l'oreille.

8.2.2 Critères de parcimonie

Le critère de parcimonie détermine les éléments qui vont être choisis à partir d'un dictionnaire redondant.

Approximation locale Une partie des méthodes cherche à trouver une bonne approximation locale du signal, afin de se rapprocher au mieux d'une approximation globale optimale. Le matching poursuit correspond à cette catégorie. Il sélectionne d'une manière itérative les atomes, afin de préserver au mieux l'énergie globale du signal [120]. En revanche, la décomposition n'est parfois pas suffisamment fidèle aux particularités locales des signaux.

Approche bayésienne D'autres recherches adoptent une approche bayésienne, en se servant des outils statistiques. Elles considèrent le caractère aléatoire des coefficients de la décomposition en accord avec une loi de distribution. Cette loi de parcimonie est choisie souvent parmi les lois de probabilité : normale, laplacienne, exponentielle, ou mélange de dirac et gaussienne. La démarche statistique intègre également des modèles dits hybrides, qui utilisent des procédés des chaînes de Markov cachées [121]. A cette catégorie appartiennent les mesures d'entropie de Renyi et Shannon [122] qui déterminent la quantité d'information du signal préservée par la décomposition.

Autres approches Une façon de chercher l'approximation parcimonieuse est fournie par la programmation convexe, qui regroupe des méthodes comme les moindres carrés, la programmation linéaire et convexe [123]. Le critère de parcimonie est remplacé dans ce cas par une fonction convexe.

8.3 Matching pursuit

L'algorithme initial de matching pursuit a été élaboré par Mallat et Zhang en 1993 [120]. Une étude de Davis élargit ce concept dans le contexte des approximations nonlinéaires adaptatives [119] [124]. Par ailleurs, Chen et Donoho développent le "basis pursuit" [125]. De nombreux algorithmes, comme par exemple "tree-based pursuit", découlent de la même idée du matching pursuit [126]. Ici on va mentionner seulement les variantes qui ont un esprit commun avec notre méthode et plus particulièrement l'idée de "ridge pursuit" introduit par Gribonval [127].

8.3.1 Algorithme

Le matching pursuit est un algorithme de décomposition itérative du signal en combinaison linéaire d'atomes, choisis à partir d'un dictionnaire $D = g_\gamma, \gamma \in \Gamma$. Dans l'algorithme initial de Mallat et Zhang [120], les atomes temps-fréquence ont la forme générale suivante

$$g_{s,u,\xi}(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\xi t}, \quad (8.3)$$

où $\gamma \in \Gamma$ est un indice, $\gamma = (s, u, \xi)$, où s est un paramètre d'échelle, u et ξ sont les paramètres temps et fréquence.

Le matching pursuit exprime une technique adaptative qui choisit à chaque itération le meilleur atome, i.e qui est le plus proche du signal x . Le critère de sélection a pour objectif de préserver au mieux l'énergie du signal.

Etape 1 : Le premier atome est choisi comme celui qui maximise la fonction $|\langle x, g_\gamma \rangle|^2$. Cette opération comprend le calcul des produits scalaires du signal x avec tous les éléments du dictionnaire, afin de trouver l'indice γ_1 du meilleur candidat,

$$\gamma_1 = \arg \max_{\gamma \in \Gamma} |\langle x, g_\gamma \rangle|^2. \quad (8.4)$$

Etape 2 : Le meilleur atome trouvé est ensuite soustrait du signal. On obtient un résidu R^1 , dont l'énergie est égale à l'énergie du signal initial moins l'énergie de l'atome enlevé.

$$R^1 x = x - \langle x, g_{\gamma_1} \rangle g_{\gamma_1}. \quad (8.5)$$

Etape 3 : D'une manière itérative, on répète la sélection du meilleur atome dans le reste du dictionnaire

$$\gamma_{m+1} = \arg \max_{\gamma \in \Gamma} |\langle R^m x, g_\gamma \rangle|^2, \quad (8.6)$$

et on actualise le résidu R^{m+1} de la $m + 1$ -ième itération

$$R^{m+1} x = R^m x - \langle R^m x, g_{\gamma_{m+1}} \rangle g_{\gamma_{m+1}}. \quad (8.7)$$

Résultat : Après M itérations, on obtient une approximation x_M du signal x à M -termes :

$$x_M = x - R^M x = \sum_{m=1}^M \langle R^{m-1} x, g_{\gamma_m} \rangle g_{\gamma_m} \quad (8.8)$$

C'est un algorithme qui donne la construction pas à pas de l'approximation, en commençant avec l'approximation initiale du signal $s^{(0)} = 0$ et un résidu $R_{(0)} = s$.

Discussion L'algorithme du matching pursuit est un pas important vers les représentations incomplètes, mais l'inconvénient principal réside dans le fait qu'il nécessite un nombre d'opérations très élevé. En effet, à chaque itération, il faut parcourir tout le dictionnaire à la recherche du meilleur atome et calculer les produits scalaire entre les atomes et le résidu. Un algorithme similaire, qui utilise un dictionnaire d'atomes temps-fréquence de Gabor, est développé indépendamment par Qian et Chen [128]. Le matching pursuit ne fournit pas la meilleure décomposition, et les atomes dans la sélection finale, ne sont pas, en général, orthogonaux. De plus, la décomposition avec un nombre fini d'itérations n'est pas garantie. Une variante est le matching pursuit orthonormal, qui applique sur la famille de sélection $\{g_{\gamma_m}\}$, l'orthogonalisation de Gram-Schmidt [129]. Le matching pursuit ainsi modifié converge plus rapidement. Cependant, l'instabilité numérique et le coût de calcul qui accompagne cette démarche est beaucoup plus important que dans l'algorithme initial [124].

8.3.2 Atomes réels

Quand on travaille avec des atomes réels

$$g_{(\gamma,\phi)} = \frac{K_{(\gamma,\phi)}}{\sqrt{s}} g\left(\frac{t-u}{s}\right) \cos(\xi t + \phi) \quad (8.9)$$

la phase, auparavant "cachée" dans le nombre complexe, apparaît d'une manière explicite. La constante $K_{(\gamma,\phi)}$ est déterminée de façon que l'atome soit unitaire $|g| = 1$. L'atome réel est relié à son complexe par l'expression

$$g_{(\gamma,\phi)} = \frac{K_{(\gamma,\phi)}}{2} (e^{i\phi} g_\gamma + e^{-i\phi} g_{\gamma-}) \quad (8.10)$$

et la constante de normalisation est

$$K_{(\gamma,\phi)} = \frac{\sqrt{2}}{\sqrt{1 + \operatorname{Re}(e^{i2\phi} \langle g_\gamma, g_{\gamma-} \rangle)}}. \quad (8.11)$$

Dans l'expression de l'atome réel apparaissent à la fois les fréquences positives et négatives. De plus l'amplitude n'est pas la même que dans le cas complexe. Par conséquent, le matching poursuit dans le cas réel et dans le cas complexe ne sont pas identiques. En effet, les atomes g_γ et $g_{\gamma-}$ ne sont pas orthogonaux entre eux. Cette situation est étudiée dans la version de Mallat et Zhang où le matching poursuit sur des atomes réels calcule le produit scalaire avec l'atome complexe et le résidu et déduit la phase à partir de la forme exponentielle du nombre complexe en termes de module et argument,

$$e^{i\phi_{arg}} = \frac{\langle R^{m-1}x, g_{\gamma_m} \rangle}{|\langle R^{m-1}x, g_\gamma \rangle|}. \quad (8.12)$$

Même après tous les efforts d'optimisation un facteur primordial de la complexité algorithmique du matching poursuit reste la taille du dictionnaire.

8.4 Extensions du matching poursuit

En s'inspirant de l'algorithme de matching poursuit, dans le but de l'améliorer, plusieurs généralisations sont obtenues. Ces variantes se proposent de réduire la complexité de calcul, mais également apportent des idées nouvelles sur les approximations parcimonieuses.

8.4.1 Atomes chirps multi-échelle

Les atomes chirps présentent une alternative intéressante au dictionnaire standard, car ils permettent l'adaptation plus souple aux structures locales du signal. Ceci est d'autant plus valable pour l'approximation plus efficace des signaux modulés en fréquence.

Définition Le dictionnaire constitué d'atomes chirp multi-échelle de Gabor comporte un indice de quatre paramètres à la place de trois, i.e $\gamma = (s, u, \xi, c)$, et les paramètres correspondent à la dilatation s , translation u , modulation ξ et enfin, la modulation du chirp c . L'atome chirp multi-échelle s'écrit

$$g_{(s,u,\xi,c)}(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i[\xi(t-u) + \frac{c}{2}(t-u)^2]}. \quad (8.13)$$

Le paramètre de chirp c traduit le fait que l'atome possède une fréquence instantanée ω_i qui n'est pas constante, mais varie linéairement dans le temps. La loi de variation trace une droite dans le plan temps-fréquence $\omega_i = \xi + ct$, où le paramètre de chirp c détermine l'orientation (la rotation) de l'atome.

Applications Les atomes chirp sont introduits pour analyser certains types de signaux nonstationnaires, dont la fréquence instantanée suit une loi linéaire. Dans la littérature, Baraniuk et Jones se servent des chirps pour construire des bases et frames orthogonales [130]. Par ailleurs, Mann et Haykin définissent la transformée en chirplettes (en anglais, "chirplet transform"). Elle représente un équivalent de la transformée de Gabor, qui utilise une gaussienne chirpée à la place de la fenêtre d'analyse standard [131][132]. L'algorithme de matching poursuit est appliqué pour la première fois sur un dictionnaire d'atomes chirps par Bultan [133]. Le fait d'introduire un quatrième paramètre, le paramètre du chirp, augmente d'une façon significative la complexité de calcul. Les applications sont limitées donc à des signaux de petite

taille et à un nombre d'itérations réduit à une dizaine. Pour réduire le temps de calcul, Bultan restreint la discrétisation du paramètre du chirp aux grandes échelles. Par conséquent, l'estimation du paramètre du chirp reste imprécise. Malgré l'intérêt de l'algorithme de Bultan, le problème avec la complexité de calcul de l'ordre de $O(MN \log N)$ (où M est le nombre d'itération et N , nombre d'éléments dans le dictionnaire) limite ses applications.

8.4.2 Ridge pursuit

Récemment, Gribonval a proposé une modification de l'algorithme de matching pursuit et réduit la complexité en $O(MN)$. Dans son approche, il utilise des atomes gaussiens, ainsi que les maxima locaux du dictionnaire de Gabor et les techniques des arêtes. Les atomes chirps forment un dictionnaire D^+ qui est une extension du dictionnaire de Gabor standard D . Le large nombre d'éléments dans D^+ ne permet pas de suivre la procédure de matching pursuit, qui parcourt à chaque itération le dictionnaire.

Idée Afin de remédier à cette restriction, Gribonval propose d'effectuer la recherche du meilleur atome chirp en deux temps [127]. Dans le premier, il sélectionne le meilleur atome de Gabor non chirpé

$$g_{(s_m^*, u_m^*, \xi_m^*)} = \arg \max_{g \in D} |\langle R^{m-1} x, g_{(s, u, \xi)} \rangle|, \quad (8.14)$$

et par la suite, il explore son voisinage dans D^+ pour ajuster la valeur du meilleur atome chirp

$$g_{(s_m, u_m, \xi_m)} = \arg \max_{g \in D^+} |\langle R^{m-1} x, g_{(s_m, u_m^*, \xi_m^*), c} \rangle|. \quad (8.15)$$

Le dictionnaire de Gabor D est complet, i.e il contient toute l'information nécessaire pour calculer les produits scalaires $\langle R^{m-1} x, g_\gamma \rangle$. Une justification de ce fait est donnée dans la continuité des résultats sur l'analyse des variations de la fréquence instantanée des signaux non-stationnaires [72] et l'extraction des ridges [71] [134]. En effet, le meilleur atome non chirpé est situé sur les ridges. Gribonval montre que le résidu $R^{m-1} x$, analysé à travers d'un atome chirp ressemble à un autre atome chirp du D^+ . Ce fait est exprimé par l'approximation

$$\langle R^{m-1} x, g_{(s_m, u_m, \xi_m, c)} \rangle \approx A(u) e^{i\Phi(u)} \langle g_{\gamma^+(u)}, g_{(s_m, u_m, \xi_m, c)} \rangle. \quad (8.16)$$

Cette approximation permet d'avoir une estimation rapide des paramètres. Par ailleurs, comme dans l'algorithme de matching pursuit utilisant des atomes réels non chirpés, on effectue l'analyse des signaux réels à l'aide des atomes de Gabor réels

$$g_{(s, u, \xi, c, \phi)} = K_{(s, u, \xi, c, \phi)} g \left(\frac{t-u}{s} \right) \times \cos \left[\xi(t-u) + \frac{c}{2}(t-u)^2 + \phi \right]. \quad (8.17)$$

Algorithme L'algorithme de ridge pursuit utilise les informations locales des maxima et les ridges (arêtes) dans un dictionnaire de chirps linéaires de Gabor. Cet algorithme comprend plusieurs étapes, dont les deux premières font partie de l'algorithme de ridge pursuit "rapide".

Etape 1 : A partir du dictionnaire initial de Gabor D , il est constitué un sous-dictionnaire D_m qui comprend les P maxima locaux. Les maxima locaux sont estimés de façon que l'atome $g_{(s, u, \xi)}$ ait un maximum local à la fois par rapport aux deux paramètres u - de temps et ξ - de fréquence. Le nombre P est fixé de manière arbitraire.

Etape 2 : Une procédure d'estimation locale du meilleur atome chirp est appliquée pour chaque atome dans D_m . Cet ensemble D_m^+ d'atomes chirps ainsi sélectionnés, est un sous-ensemble du dictionnaire des atomes chirps D^+ .

Etape 3 : Pour chaque atome complexe de Gabor est calculé le produit scalaire $\langle R^{m-1} x, g_{(s, u, \xi)} \rangle$.

Etape 4 : Il est calculé le produit scalaire du résidu avec l'atome réel, ainsi que la phase optimale correspondant à l'atome réel. La position (s_m^*, u_m^*, ξ_m^*) du meilleur atome réel de Gabor est sélectionnée.

Etape 5 : L'estimation locale des paramètres s_m et c_m se fait à l'aide d'une interpolation parabolique.

Etape 6 : Il est calculé le produit scalaire $\langle R^{m-1} x, g_{(s_m, u_m^*, \xi_m^*, c_m)} \rangle$ et est déterminé le meilleur atome réel chirpé $g_{(s_m, u_m^*, \xi_m^*, c_m, \phi_m)}$. Le résidu est mis à jour.

8.5 Exemples

Les deux exemples suivants illustrent la décomposition du signal d'un quart de cycle moteur et d'un cycle moteur par l'algorithme du matching poursuit. J'ai utilisé l'interface Guimauve, créée par F. Brachere [135], qui se base sur l'implémentation du matching poursuit de LastWave par E. Bacry [136].

Exemple 1 : Un quart de cycle moteur Un exemple de la décomposition fournie par le matching poursuit sur un quart de cycle moteur est présenté sur la **Fig. 8.1**. Ce résultat est obtenu à l'aide du logiciel Guimauve, en utilisant les paramètres par défaut. On peut apercevoir les structures "verticales" et horizontales" dans le signal décomposé. La décomposition utilise 100 atomes.

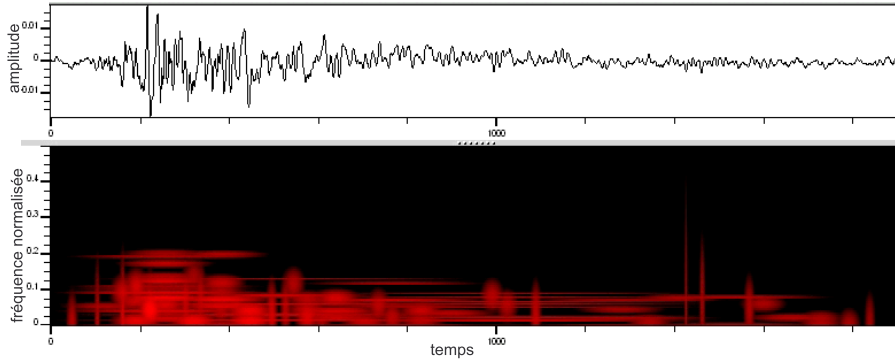


FIG. 8.1 – Décomposition d'un quart de cycle moteur à l'aide du matching poursuit (100 atomes)

Exemple 2 : Un cycle moteur L'exemple de la décomposition matching poursuit d'un cycle moteur sur 300 atomes est présentée sur la **Fig. 8.2**. L'excellente approximation du signal original par le signal reconstruit d'après la décomposition du cycle moteur témoigne des avantages de cette méthode très élaborée : le matching poursuit (**Fig. 8.3**).

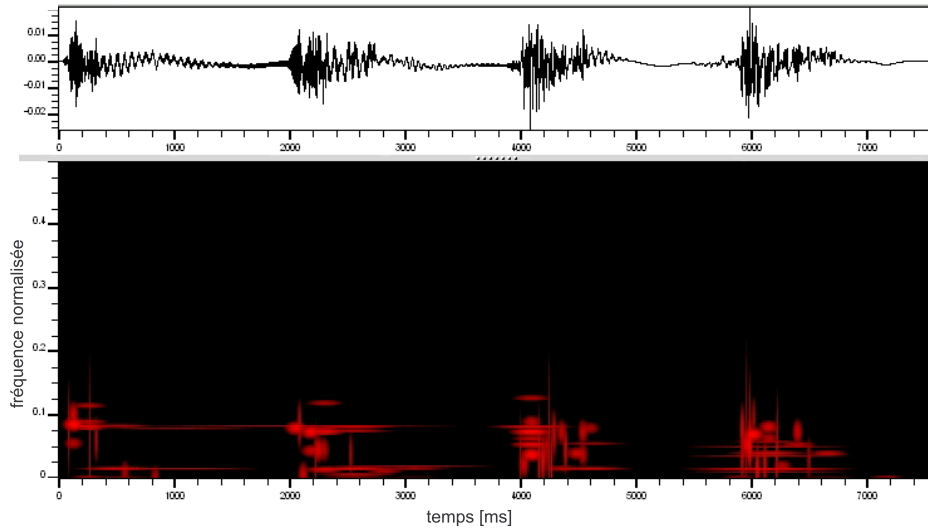


FIG. 8.2 – Décomposition d'un cycle moteur à l'aide du matching poursuit (300 atomes)

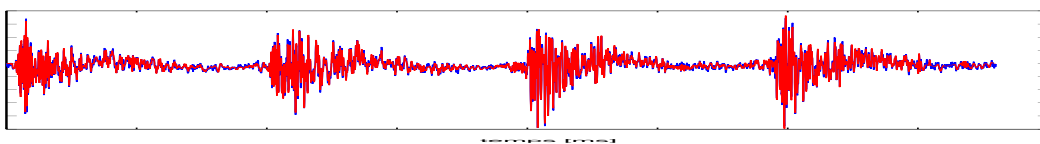


FIG. 8.3 – Signal original (en bleu) et signal reconstruit (en rouge) par le matching poursuit (300 atomes)

Conclusion sur les décompositions atomiques

Dans ce chapitre je me suis proposée de situer l'approche des décompositions atomiques du signal. L'idée directrice de ces décompositions vient de Gabor et son "quantum de son", la petite unité de l'atome temps-fréquence qui constitue le signal. Les décompositions atomiques tracent ainsi le lien avec un autre aspect de cette idée, la synthèse granulaire. On peut visualiser cette relation comme la suite "analyse-décomposition-synthèse". D'une manière générale, les décompositions sur une base orthogonale ont l'avantage de donner une solution unique du problème. Cependant, l'orthogonalité reste parfois "rigide" envers la nature des signaux. La recherche d'une décomposition optimale dans un dictionnaire d'atomes est un problème qui n'a pas de solution. Un algorithme pilier des décompositions atomiques est le matching pursuit. C'est une méthode qui possède de nombreuses variantes.

Orientation de la thèse : La nouvelle méthode d'analyse- synthèse présente le signal comme une somme d'atomes chirpés à nombre restreint. Elle fournit donc une décomposition très parcimonieuse du signal et en plus, ne satisfait pas les propriétés d'orthogonalité entre les atomes. Cette décomposition s'adapte localement au signal. Sur ce point, elle diffère de l'algorithme standard du matching pursuit qui cherche une décomposition orthogonale globale du signal, mais en revanche s'approche du ridge pursuit qui est davantage basé sur les propriétés locales du signal via les maxima locaux des atomes dans le dictionnaire. Par ailleurs, il me semble important de souligner que l'orthogonalité est une exigence mathématique et il n'est pas évident d'estimer jusqu'à quel point elle transcrit les caractéristiques perceptives du signal.

Synthèse sonore

Ce chapitre discute des méthodes de synthèse sonore utilisées pour les bruits moteur. Le but n'étant pas de faire une classification exhaustive, j'ai choisi de présenter trois méthodes de synthèse : la synthèse additive, la synthèse par modèles physiques et la synthèse granulaire. Les plus importantes applications pour les bruits moteur sont faites par synthèse additive ou par modèles physiques, alors que la synthèse granulaire n'a pas été largement utilisée pour les bruits moteur. Toutefois, afin de situer notre méthode, je trouve indispensable de décrire la synthèse granulaire non seulement dans le contexte de l'approche d'analyse-synthèse, mais également dans la démarche des compositeurs de la musique électro-acoustique. Leurs recherches mettent en évidence des phénomènes perceptifs très intéressants qui apparaissent lors de la modification des paramètres des grains. La description de ces phénomènes facilitera l'introduction de la théorie des scènes auditives, que je ferai dans la troisième partie de la thèse, consacrée à l'étude perceptive des bruits moteur.

9.1 Classification des méthodes

La classification des méthodes de synthèse la plus répandue distingue deux grandes groupes de méthodes : méthodes dites "de signal" et méthodes physiques. Les méthodes de signal ne s'intéressent pas à la source physique émettrice du son, alors que les méthodes physiques se servent des caractéristiques mécaniques de la source afin de reproduire son comportement vibro-acoustique. Enfin, il existe des méthodes dites "hybrides", combinant les approches des méthodes de signal et des méthodes physiques (**Fig. 9.1**).

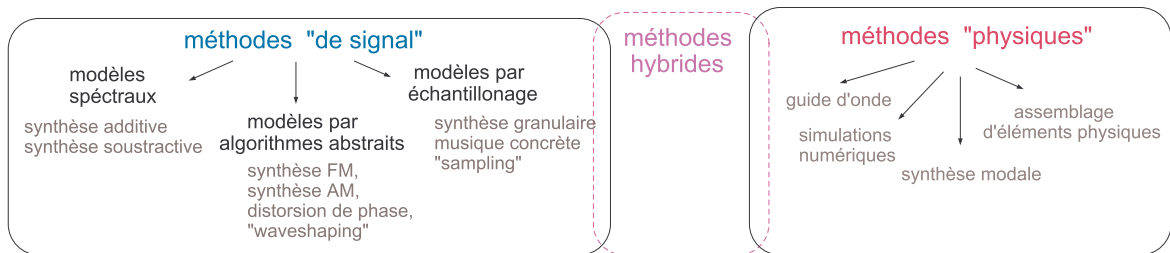


FIG. 9.1 – Classification des méthodes de synthèse

Afin de situer la synthèse additive, la synthèse par modèles physiques et la synthèse granulaire, on va brièvement parcourir les différents types de méthodes de synthèse. Cette présentation est inspirée en grande partie par la classification des méthodes de synthèse de J. Smith [137]

9.1.1 Méthodes "de signal"

Les méthodes dites "de signal" considèrent les sons d'un point de vue de traitement du signal, en s'intéressant aux paramètres comme la forme d'onde, le spectre, la phase ou l'amplitude des sons. Les méthodes les plus populaires sont la synthèse additive, la synthèse soustractive et la synthèse FM. En général, les méthodes de synthèse "de signal" ont la faculté de créer des sonorités très intéressantes, car inexistantes dans la réalité physique. En revanche, les modifications apportées dans les sons de synthèse ne sont plus reliées aux paramètres physiques de la source, comme c'est le cas des sons obtenus par modélisation physique.

Modèles spectraux Les modèles spectraux sont orientés vers la reproduction ou la modification du spectre des sons. La synthèse additive, qu'on présente par la suite, construit le son en additionnant des ondes de type sinusoïdal. D'une manière générale, c'est une méthode très adaptée pour les sons de nature tonale. La synthèse soustractive utilise un procédé de soustraction (filtrage) d'une partie du spectre d'un signal riche en composantes, le plus souvent un bruit blanc. L'avantage des modèles spectraux réside dans les algorithmes automatiques d'extraction des paramètres à partir d'un signal original. Leur inconvénient

principal reste la gestion d'un grand nombre de paramètres de synthèse.

Modèles par algorithmes abstraits Les modèles dans cette catégorie utilisent des algorithmes divers pour la génération des objets sonores. La synthèse la plus élaborée est la synthèse par modulation de fréquence, ou encore synthèse FM, découverte par Chowning [138]. Elle représente un son comme la modulation en fréquence d'une onde porteuse par une onde modulante. Parmi les autres méthodes de cette catégorie on peut citer la synthèse par distorsion non linéaire ("waveshaping"), la synthèse par distorsion de phase, la synthèse par modulation d'amplitude (MA), la synthèse vectorielle. Ces méthodes sont utilisées principalement dans la musique électroacoustique. Leur avantage réside dans le fait qu'en changeant peu de paramètres, des sonorités très différentes peuvent être obtenues. En revanche, la resynthèse à partir de sons originaux physiques peut s'avérer insuffisamment réaliste.

Modèles par échantillonnage Historiquement, cette forme de synthèse est introduite par Schaeffer dans ses expérimentations sur la musique concrète [139]. Elle consiste dans la modification ou le pilotage en temps réel des enregistrements sonores. Ce type de méthodes comprend également la synthèse par échantillonnage ("sampling") ou encore la synthèse granulaire. Toutefois, comme on le verra plus tard, la synthèse granulaire peut être affiliée à plusieurs catégories de méthodes de synthèse. L'échantillonnage est utilisé dans les instruments de musique numériques actuels et dans la création sonore pour les jeux vidéo. Toutefois, il est discutable de qualifier cette méthode de méthode de synthèse.

9.1.2 Méthodes physiques

Les méthodes dites "physiques" se réfèrent à une source physique afin de reproduire le son. Habituellement, dans la catégorie des méthodes de synthèse par modèles physiques sont considérées des techniques comme les simulations numériques, la synthèse modale, les modèles guide d'onde numérique et certains modèles dits hybrides.

Simulations numériques Les simulations numériques cherchent à résoudre numériquement des équations physiques qui décrivent le comportement d'un système mécano-acoustique. Ceci revient à chercher dans le domaine discret une approximation de la solution des équations. Le plus souvent, la modélisation dans le domaine acoustique se fait à l'aide des équations aux dérivées partielles, comportant des dérivées spatiales et temporelles. Des exemples de simulations numériques sont le calcul par différences finies, par éléments finis (FEM) ou éléments frontières (BEM). L'avantage de cette technique est sa précision dans la description des phénomènes physiques. L'inconvénient principal est que l'augmentation de la précision entraîne souvent une complexité de calcul importante.

Modèle de guide d'onde numérique Tandis que les simulations numériques cherchent une solution approchée, les modèles de guide d'onde numérique se servent directement des solutions connues des équations physiques. Le modèle le plus simple de guide d'onde décrit la propagation du premier mode transverse d'une onde acoustique dans un tube monodimensionnel. Il est également possible de modéliser des systèmes physiques plus complexes, avec ou sans perte. Les modèles de guide d'onde numérique sont développés pour la synthèse des sons des instruments de musique par J. Smith [140].

Synthèse modale La synthèse modale est orientée vers l'étude des petites oscillations des structures linéaires faiblement amorties. Cette méthode utilise l'information de la masse, l'amortissement et la raideur de la structure, ainsi que des différentes forces d'inertie, de dissipation et des forces extérieures qui excitent la structure. La synthèse modale fournit les modes propres de vibrations du système vibro-acoustique. La première application de la synthèse modale sont les instruments de musique, mais cette méthode trouve également en usage dans d'autres domaines, y compris l'automobile.

Modèles virtuels Parmi les autres modèles physiques figurent des modèles virtuels ou encore des modèles particuliers, comme Cordis-Anima, qui consistent à assembler des éléments physiques. Cordis-Anima permet de modéliser un processus physique à l'aide d'un langage qui utilise des objets virtuels, correspondants à un comportement mécanique comme par exemple des systèmes de type masse-ressort-amortisseur [141].

9.1.3 Concept d'analyse-synthèse

La classification exposée ci-dessus représente juste une des visions possibles des méthodes de synthèse. En parlant du procédé d'analyse-synthèse, on peut considérer un autre point de vue, en distinguant deux approches de la synthèse sonore : synthèse à partir d'un signal original et synthèse sans un signal de référence.

Synthèse à partir d'un signal de référence La première approche part d'un signal original, passe par un procédé d'estimation des paramètres de synthèse, et enfin, reconstruit le signal (**Fig. 9.2**). L'efficacité du procédé d'estimation est jugée en premier selon la ressemblance perceptive entre le signal de synthèse et le signal original. Pour deux procédés différents, qui fournissent un résultat perceptif presque identique, c'est celui qui sera d'une moindre complexité de calcul qui sera préféré.

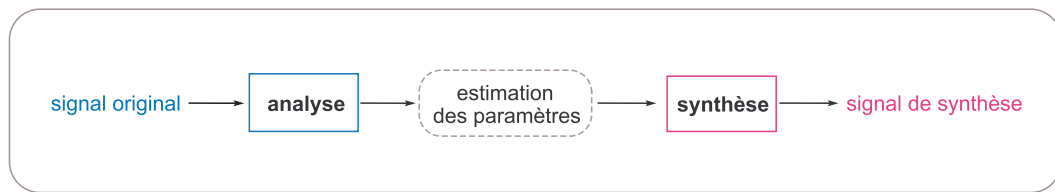


FIG. 9.2 – La démarche globale d'analyse-synthèse des signaux sonores

Les paramètres à estimer sont déterminés en fonction de la méthode d'analyse-synthèse choisie. Chaque méthode possède sa façon de paramétrer le signal. On verra par la suite que pour la synthèse additive, qui présente le signal comme une somme de sinusôides, les paramètres de synthèse à estimer sont les 3 paramètres de chaque sinusôide : l'amplitude, la phase et la fréquence instantanée. De même, pour la synthèse granulaire, les paramètres à estimer sont le temps d'apparition des grains, ainsi que leur fréquence, durée, enveloppe et amplitude. La synthèse par modèle physique modélise le comportement acoustique des objets sonores et utilise comme paramètres les propriétés mécaniques des matériaux qui les constituent. La démarche d'analyse-synthèse donne la possibilité de modifier les valeurs des paramètres estimés et dans le cas de la synthèse par modèle physique, même de les relier à une réalité physique.

Synthèse sans signal de référence La deuxième approche de la synthèse sonore ne se réfère ni à un signal original, ni à une source sonore. Elle part d'une sorte de "tabula rasa" ¹, sans "a priori" sur le son, et essaie de sculpter des différentes sonorités. Certaines sonorités peuvent se rapprocher des sons qui nous sont communs, comme par exemple les instruments de musique. Cette deuxième méthode de synthèse possède également des paramètres de contrôle, comme par exemple la modulation de fréquence (FM), qui est gérée à l'aide des paramètres de fréquence de la porteuse et la modulante, ainsi que l'indice de modulation. Cependant, cette démarche ne recourt pas à une estimation de paramètres proprement dite. Par conséquent, il est souvent difficile de bénéficier d'une extraction automatique des paramètres de synthèse pour un son spécifique, comme ça se fait dans le procédé d'analyse-synthèse. Les paramètres doivent alors être ajustés expérimentalement. Toutefois, ceci ne nuit point au réalisme sonore qui peut être atteint à l'aide des méthodes qui n'analysent pas un signal de référence. Ceci est le cas de la synthèse FM qui reproduit assez fidèlement les timbres des instruments de musique. Parfois, comme dans la synthèse granulaire, c'est justement le fait de ne pas avoir de signal de référence qui est apprécié, car il est possible de créer, en partant de rien, des structures sonores inouïes.

Remarque Cette division en deux catégories selon l'utilisation ou pas d'un signal de référence, ne peut pas servir de classification des méthodes de synthèse. En effet, en fonction des objectifs de synthèse, une méthode peut appartenir et à l'une et à l'autre catégorie. Par exemple, la synthèse granulaire et la synthèse additive peuvent être utilisées aussi bien comme dans le procédé d'analyse-synthèse, que dans la synthèse qui construit un signal à partir de rien. Evidemment, la synthèse granulaire que j'ai exploitée pour les bruits au ralenti, est attribuée à la première catégorie, celle de l'analyse-synthèse.

¹"tabula rasa" (littéralement "table rase") concept philosophique selon lequel l'esprit humain naîtrait vierge et serait formé par l'expérience vécue.

9.2 Synthèse additive

La synthèse additive est probablement la méthode de synthèse la plus populaire. Cette popularité est due en grande partie à son universalité. En effet, grâce aux algorithmes d'extraction de paramètres, la synthèse additive peut être appliquée pour une large palette de sons harmoniques et non harmoniques, aussi bien dans un but de création sonore que pour la resynthèse des sons du monde réel. Cette section définit la démarche de la synthèse additive en vue de l'extraction des paramètres du signal et donne un exemple d'application pour les bruits moteur.

9.2.1 Démarche

L'appellation de la synthèse additive suggère qu'elle fonctionne par addition des éléments. Dans le domaine de la vision, elle consiste à combiner les couleurs dites de base : le rouge, le vert et le bleu, et à additionner leurs lumières afin d'obtenir une multitude de couleurs. L'addition des couleurs de base donne le blanc. Dans l'audition, les sons complexes sont obtenus par sommation des ondes simples, sinusoïdales. Le bruit blanc contient toutes les fréquences du spectre.

Définition Fourier découvre que tout mouvement périodique complexe se décompose en une somme d'ondes sinusoïdales, appelées harmoniques, dont les fréquences sont des multiples entiers de la fréquence la plus basse, appelée fondamentale. Ce concept est à la base de la synthèse additive qui suppose que les sons harmoniques peuvent être considérés comme la somme de sinusoïdes, dont l'amplitude et la fréquence évoluent dans le temps. Le signal est modélisé par la formule

$$s(t) = \sum_{k=1}^N A_k(t) \cdot \cos[\Theta_k(t)] \quad \text{avec} \quad \Theta_k(t) = \int_0^t \omega_k(\tau) d\tau + \phi_k, \quad (9.1)$$

où $A_k(t)$, $\Theta_k(t)$ et $\omega_k(t)$ sont l'amplitude, la phase et la fréquence instantanées de la k -ème composante. Pour les sons harmoniques, la première sinusoïde dans la série ($k = 1$) est la fréquence fondamentale, alors que les sinusoïdes de rang plus élevé représentent les harmoniques. Le modèle de la synthèse additive (eq. 9.1) est à l'origine des algorithmes d'extraction des paramètres, décrits peu après.

Applications Historiquement, la première manifestation de la synthèse additive à "action manuelle" date du Moyen Âge. Elle est associée au plus grand instrument, l'orgue. L'orgue est constitué d'un ensemble de tuyaux où l'air, passant par chaque tuyau, crée une seule note, une harmonique. Les sons plus complexes et riches en harmoniques sont obtenus en actionnant un plus grand nombre de tuyaux ouverts, la combinaison étant déterminée par un registre donnant un timbre spécifique [142].

Le premier véritable instrument de synthèse additive musicale est le Telharmonium créé par Cahill [143]. Le son est produit par l'action électromécanique des crans d'une roue phonique générant des signaux électriques, dont la fréquence dépend de la vitesse de rotation et du nombre de crans. Un perfectionnement de ce principe est à l'origine de l'orgue Hammond, l'instrument électronique de la synthèse additive de référence [144]. D'un point de vue électronique, la synthèse additive est implémentée comme une somme d'oscillateurs qui génèrent les différents partiels du son. Chaque partiel est déterminé par son amplitude et sa fréquence (**Fig. 9.3**).

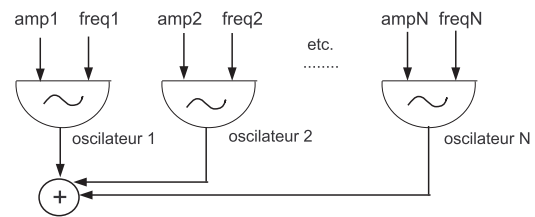


Fig. 9.3 – Diagramme d'un synthétiseur additif

Avantages et limites de la méthode La synthèse additive permet facilement la modélisation et la modification des composantes du signal. Les paramètres de la synthèse additive (fréquence, amplitude, phase) sont très intuitifs et ils peuvent être modifiés indépendamment. Il est également possible d'implémenter des indicateurs perceptifs comme la rugosité, l'inharmonicité. En effet, le modèle additif est un modèle spectral de synthèse qui est basé sur la représentation spectrale des sons dans l'oreille humaine. Or, la plupart des modèles psychoacoustiques sont établis en accord de ce modèle additif. En revanche, la gestion d'un grand nombre de partiels peut s'avérer difficile. Par ailleurs, la synthèse additive semble moins bien adaptée pour la synthèse des sons impulsifs.

9.2.2 Extraction des paramètres

La synthèse additive présente les sons comme une somme de sinusôides dont l'amplitude et la fréquence évoluent dans le temps. Trois modèles d'extraction tracent l'évolution de la synthèse additive : le modèle "sinusoïdal" de McAulay et Quatieri, le modèle "sinusoïdes + bruit" de Serra et le modèle "sinusoïdes + transitoire + bruit" de Verma.

Modèle sinusoïdal Ce modèle le plus simple suit la présentation d'un son harmonique comme une somme de sinusôides (Eq. 9.1) et est introduit par McAulay et Quatieri pour la synthèse de la parole [145]. Après l'application de la TFCT sur le signal, les fréquences et les amplitudes sont extraites à partir des pics (i.e. les maxima locaux) dans le spectre d'amplitudes. Les phases sont calculées à partir du spectre de phase, aux mêmes valeurs que les pics spectraux (**Fig. 9.4**).



FIG. 9.4 – Schéma d'analyse-synthèse du modèle sinusoïdal

La synthèse se fait fenêtre par fenêtre, après l'interpolation des amplitudes, le déroulement et l'interpolation de la phase. Les sinusôides sont ensuite additionnées pour donner le signal de synthèse.

Modèle sinusoïdes + bruit Afin de tenir compte de la réalité physique des sons naturels, un autre modèle de synthèse additive est proposé par Serra [146]. Il considère le signal comme composé de deux parties : déterministe et stochastique, i.e comme

$$s(t) = det(t) + res(t) = \sum_{k=1}^N A_k(t).cos[\Theta_k(t)] + res(t) \quad (9.2)$$

La partie déterministe $det(t)$ modélise le contenu harmonique du signal comme une somme de sinusôides. La partie résiduelle $res(t)$ est considérée comme un bruit aléatoire. L'extraction de la partie déterministe passe par la détection des pics, par l'estimation de la fréquence fondamentale et par l'estimation des amplitudes et des phases des harmoniques. La resynthèse de la partie déterministe et sa soustraction dans le domaine temporel au signal original donne le résidu. Le résidu est modélisé par la suite comme un bruit blanc filtré dont les coefficients sont identifiés par prédiction linéaire (**Fig. 9.5**).

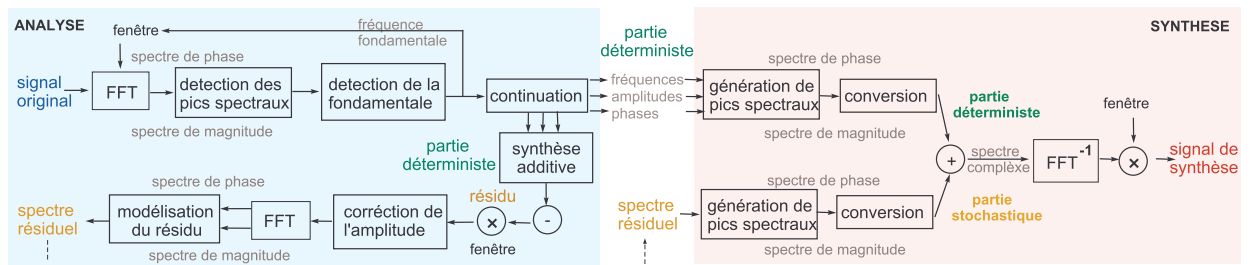


FIG. 9.5 – Schéma d'analyse du modèle sinusoïdal+bruit

La synthèse s'effectue séparément sur les parties déterministe et résiduelle et donne ainsi la possibilité de les modifier. Ce modèle connaît plusieurs implémentations : l'algorithme initial, appelé SMS (Spectral modeling synthesis) par Serra [147], Additive de l'Ircam [148] et SAS de Marchand [149].

Modèle sinusoïdes + transitoires + bruit Ce modèle de Verma apporte une amélioration importante de la synthèse additive, permettant le traitement des transitoires [150]. Il considère le signal comme composé de trois parties : sinusoïdale, transitoire et du bruit

$$s(t) = \sum_{k=1}^N A_k(t).cos[\Theta_k(t)] + r_t(t) + r_b(t), \quad (9.3)$$

où $r_t(t)$ concerne la partie transitoire et $r_b(t)$ le bruit résiduel.

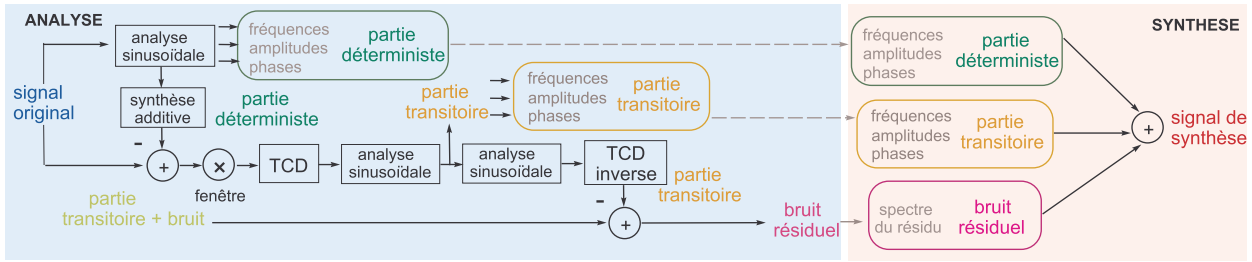


FIG. 9.6 – Schéma d'analyse-synthèse du modèle sinusoïdal + transitoire + bruit

L'algorithme extrait en premier les composantes sinusoïdales, effectue leur resynthèse additive et les soustrait du signal original. Ainsi est obtenue une première partie qui contient les éventuels transitoires et du bruit résiduel (Fig. 9.6). L'extraction des transitoires est basée sur la dualité entre sinusoïdes et transitoires, au sens où une sinusoïde dans le domaine temporel correspond à un pic fréquentiel et l'inverse pour le transitoire. Cette correspondance est faite à l'aide de la transformée en cosinus discret (TCD). Une soustraction de la partie transitoire détermine le bruit résiduel. La synthèse est effectuée séparément sur les trois parties du signal. Le modèle de Verma est mieux adapté pour les sons réels, qui peuvent contenir des transitoires en même temps qu'un spectre harmonique, par exemple l'attaque d'une note musicale. Toutefois, ce modèle n'est pas spécialement dédié aux sons impulsifs.

9.2.3 Synthèse additive des bruits moteur

La synthèse additive est la méthode la plus utilisée pour la synthèse des bruits moteur. Sa pertinence est justifiable, car le spectre des bruits moteur est principalement constitué des harmoniques multiples à la fondamentale moteur. En revanche, en tant que modèle spectral, il n'offre pas l'équivalence immédiate avec les événements temporels dans le cycle moteur. Cette section présente une compilation d'idées de l'approche additive pour la construction d'un simulateur sonore des bruits moteur.

Schéma d'un simulateur des bruits automobile La réalisation d'un simulateur utilisant la synthèse additive se décompose en 3 étapes essentielles : synthèse du bruit moteur, synthèse du bruit de roulement et du bruit aérodynamique et le contrôle dynamique des situations de conduite (Fig. 9.7).

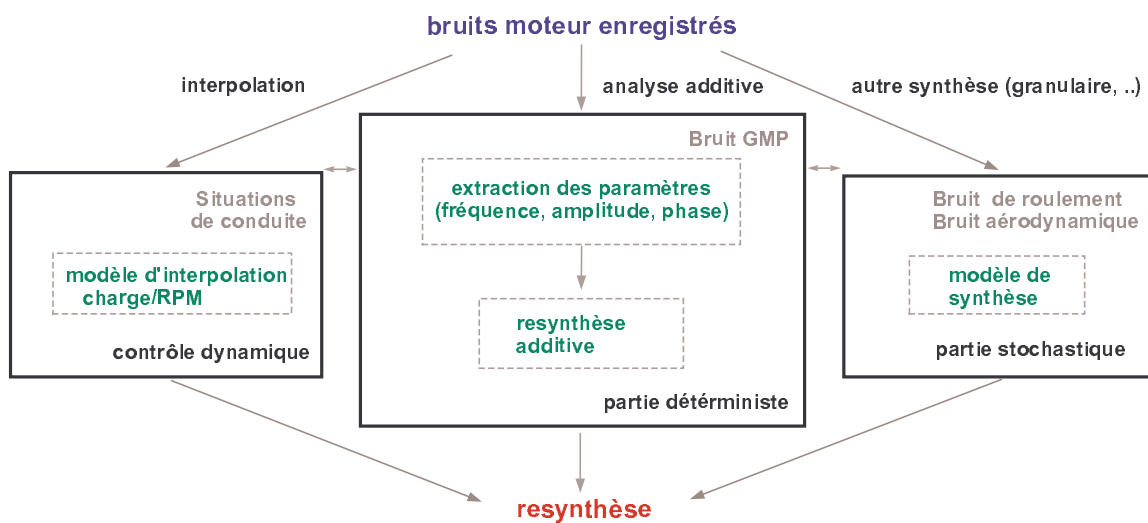


FIG. 9.7 – Schéma d'un simulateur de bruit moteur à la base de la synthèse additive

La première étape est liée à la modélisation du bruit du groupe motopropulseur (GMP). Elle consiste à extraire l'information de la fondamentale, des harmoniques et des éventuelles demi-harmoniques du bruit moteur, ainsi que leurs amplitudes et phases. Ces paramètres vont être utilisés pour la synthèse additive

de la partie déterministe du bruit moteur. Globalement, on considère comme la partie déterministe le bruit émis par le GMP, car il a un fonctionnement qui prédispose l'apparition d'harmoniques multiples à la fondamentale. En revanche, les deux autres contributions comme le bruit de roulement et le bruit aérodynamique n'ont pas forcément une nature harmonique.

La deuxième étape de la synthèse des bruits automobile concerne le bruit de roulement et le bruit aérodynamique. Ces bruits ont davantage une structure stochastique. Par conséquent, ils peuvent être vus comme le bruit résiduel dans le modèle "sinusoïdal + bruit" de Serra [151]. La modélisation de ces deux autres contributions passe par une analyse des bruits de roulement et d'écoulement enregistrés et se fait à l'aide d'un autre modèle, par exemple de synthèse granulaire ou de synthèse soustractive, qui sont mieux adaptés à la synthèse des bruits. Dans le développement du simulateur HARTIS par Richard [152] pour la synthèse de ces contributions le modèle de synthèse granulaire "Smooth overlap granular synthesis" est appliqué [153]. Par la suite, les bruits ainsi modélisés sont ajoutés au bruit final de synthèse. Ces deux premières étapes appliquent les algorithmes d'analyse-synthèse additive comme présenté dans la section précédente.

Pilotage dynamique La troisième étape dans la construction d'un simulateur temps-réel contient un pilotage dynamique des situations de conduite. Cette étape considère une carte d'accélération-décélération en fonction de l'état de la pédale d'accélération (la charge) et le régime moteur. En règle générale, les données de 25 %, 50 %, 75 % et 100 % de charge sont prises en compte, pour quelques régimes moteur fixes. Ensuite, les valeurs intermédiaires sont obtenues par interpolation linéaire entre des frames m_i du signal (**Fig. 9.8**). Pour la fondamentale moteur f_1 , l'interpolation de la valeur intermédiaire p entre deux positions mesurées de la pédale p_1 et p_2 est exprimée comme

$$f_1^p = f_1^{p_1}(m_1) + (w_{p_2}(m_2) - w_{p_1}(m_1)) \frac{p - p_1}{p_2 - p_1} \quad (9.4)$$

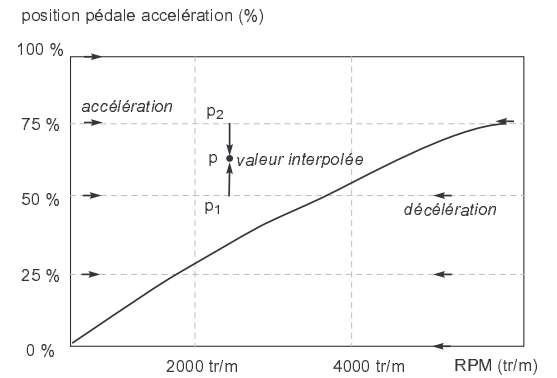


Fig. 9.8 – Interpolation des valeurs intermédiaires pour le contrôle dynamique du bruit de moteur

Toutes les valeurs intermédiaires des paramètres du modèle additif sont calculées à l'aide de cette interpolation.

Algorithme de simulation Le pilotage dynamique sous-entend un calcul d'un cycle à l'autre en fonction de l'état de la pédale d'accélération. L'algorithme de simulation peut être résumé de la manière suivante. D'abord, le régime moteur est initialisé. Par ailleurs, le régime moteur détermine la durée d'un cycle moteur qui est comprise entre 20 et 50 ms. A chaque itération, au début du cycle, la valeur de la pédale d'accélération p est prise en compte. Ensuite sont retrouvés les sons enregistrés qui correspondent aux valeurs p_1 et p_2 à partir desquelles va être interpolée la valeur p . Les trames temporelles correspondantes m_1 et m_2 vont servir pour effectuer la resynthèse sonore selon le modèle additif. Avant de passer à l'itération suivante, il est calculé, à l'aide de la fondamentale f_1 , le régime moteur pour la frame suivante. Ce schéma de pilotage dynamique et cet algorithme de simulation de bruit moteur sont proposés par Feldbauer [154].

Avantage et limite de la synthèse additive pour les bruits moteur La synthèse additive permet la conception la plus aboutie d'un simulateur de bruit moteur. Si cette approche fonctionne très bien pour les bruits à haut régime, elle ne peut pas s'adapter à un régime au ralenti qui n'a pas la nature continue, sinusoïdale, mais fortement impulsive. Le modèle additif permet la simulation temps réel, car les sinusoïdes sont faciles à implémenter. Toutefois, il est nécessaire d'estimer un grand nombre d'harmoniques pour arriver à simuler correctement le bruit moteur. Les modifications sont apportées sur l'amplitude et/ou la phase de chacune des harmoniques (ou les "demi"-harmoniques) moteur. Toutefois, la synthèse additive prouve sa pertinence perceptive sur les bruits de moteur à un régime moyen et élevé. C'est probablement le modèle le plus adapté pour un pilotage temps réel.

9.3 Synthèse par modèles physiques

La synthèse sonore par modèle physique consiste à produire un son à partir de la description du comportement physique d'un objet. Habituellement, dans la catégorie des méthodes de synthèse par modèles physique sont considérées des techniques comme les simulations numériques, la synthèse modale, les modèles guide d'onde et certains modèles dits hybrides. Comme toute classification, celle-ci est aussi très relative, car aujourd'hui il devient naturel d'associer les connaissances du domaine physique et acoustique, des outils de traitement du signal et les avancées dans la perception des sons. Les modèles physiques ont été développés d'abord pour l'étude des instruments de musique. C'est pourquoi j'ai choisi de présenter dans cette section un lien intéressant entre l'acoustique des instruments de musique et l'automobile. Par la suite, une approche de la synthèse par modèle physique appliquée aux bruits moteur est exposée.

9.3.1 Instruments de musique et automobile

Les premières applications de la synthèse modale, comme on a vu aussi dans le cas de la synthèse additive, ont été faites sur les instruments de musique. En effet, l'élaboration des instruments de musique a suscité des découvertes sur plusieurs phénomènes acoustiques et vibratoires. La découverte a été tout d'abord le fruit des années d'expériences et des secrets bien gardés dans la fabrication artisanale. Les études modernes mettent au clair les particularités des modes vibro-acoustiques des objets comme cordes, chevalets, caisses de résonances, tuyaux, cloches, plaques, membranes.


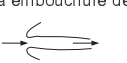
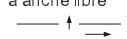
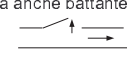
	type d'instrument	exemple	excitation	transmission/ amplification	analogie automobile
VENT	à embouchure de flute 	flûte, sirène, sifflet	bec aérodynamique	tuyau	admission d'air
	à embouchure de cor 	cor, trompette, trombone	anche = levres musicien	tuyau	échappement
	à anche libre 	orgue, accordéon, harmonica	anche	tuyau	ventilation
	à anche battante 	orgue, clarinette, haubois, sax	anche	tuyau	toit ouvrant, bruit aéro
CORDES	corde griffée corde frappée corde pincée corde frottée	clavecin, clavicord piano, guitare, harpe, luth violon, contrebasse	griffe marteau, doigt archet	table d'harmonie table d'harmonie table ame, caisse	tout système mécanique déformable: moteur, boîte, structure, carrosserie
	membrane tendue plaque	tambour, cymbale, bongo vibraphone, xylophone triangle, cloche castagnette	baguette, doigts, marteau	peau, caisse plaquettes bois, acier bois, acier, bronze	claquement piston pignons percussion de train grailonnement boîte
PERCUSSION					
FROTTEMENT					
ELECTRO					
	électronique	synthétiseur, orgue	générateur électronique	ampli	anti-bruit électronique

FIG. 9.9 – Analogie entre les phénomènes vibro-acoustiques des instruments de musique et les organes automobile

Au premier abord il paraît étonnant qu'on puisse retrouver une majorité de ces phénomènes dans l'acoustique automobile. L'observation plus fine de sa structure et son fonctionnement montre que l'automobile est une combinaison très complexes de plaques, tuyaux et d'autres éléments rotatifs. Un aperçu sur les analogies des instruments de musique et des organes automobile est donné sur la **Fig.9.9**, grandement inspiré de [155]. Cette analogie permet de situer mieux une approche par modèles physiques dans le contexte automobile.

9.3.2 Modélisation physique des bruits de moteur

Une étude récente de Guyader [156] développe la synthèse sonore des bruits moteur, basée sur la modélisation physique. Plus précisément, elle définit un processus de synthèse sonore temporelle qui utilise comme entrée des paramètres géométriques et le fonctionnement du GMP.

Processus Chaque étape du processus correspond à une réalité physique dans le moteur. La première étape consiste à générer les énergies primaires à l'origine du bruit moteur. Le signal ainsi obtenu représente l'excitation à l'entrée d'un banc de filtres, correspondant aux fonctions de transfert entre la pression cylindre et les vibrations de la structure. La dernière étape permet d'accéder à la pression acoustique rayonnée par la surface (**Fig. 12.15**). Les points clés de cette méthode sont les suivants [157].

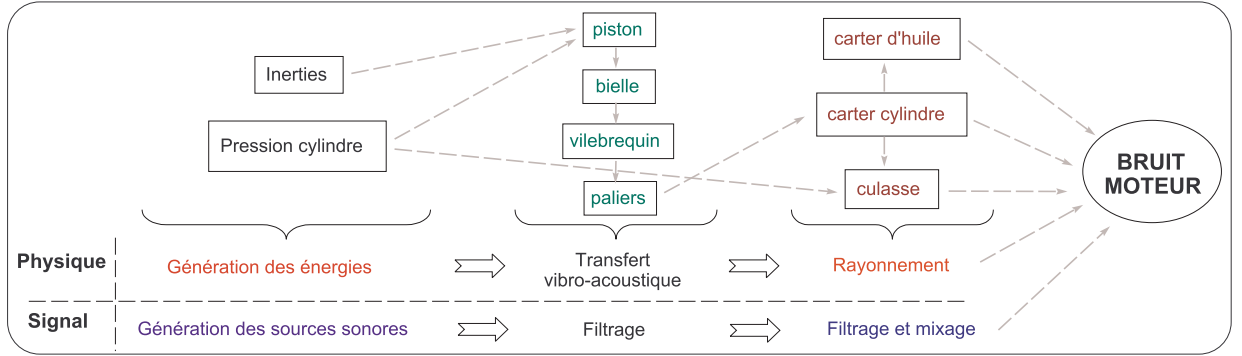


FIG. 9.10 – Schéma de synthèse sonore à partir de modélisation physique (voir [157])

Pression cylindre La première étape comprend l'estimation de la pression dans le cylindre lors des quatre temps du cycle moteur. Le point de départ du calcul est le volume $V(\theta)$ dans le cylindre, qui est calculé en fonction de l'angle vilebrequin θ

$$V(\theta) = \frac{\pi \cdot d^2}{4} \left[(L + R) - R \cdot \cos(\theta) + L \sqrt{1 - \frac{\sin^2(\theta)}{\lambda^2}} \right] + V_{pm}. \quad (9.5)$$

Dans la formule (9.5) entrent en jeu les paramètres géométriques du cylindre comme le diamètre d , ainsi que λ , qui est le rapport entre la longueur de la bielle L et le rayon manivelle R . Le volume au point mort est noté avec V_{pm} . Ensuite, la pression dans les cylindres est calculée pour chaque temps moteur. Pour la phase d'échappement et d'admission, la pression cylindre est estimée comme

$$P(\theta) = \frac{P(\theta - 1) \cdot V(\theta - 1)}{V(\theta)} + \frac{dbi(\theta - 1) \cdot V(\theta - 1)}{V(\theta)}. \quad (9.6)$$

Dans cette expression, la quantité dbi représente le débit instantané aux soupapes. Par ailleurs, il faut tenir compte de la fréquence d'échantillonnage f_e , en multipliant le débit par le pas de calcul $= \frac{RPM \cdot 360}{60 \cdot f_e}$.

La pression cylindre de la phase de compression est estimée en posant des valeurs initiales P_{in} et V_{in} de la pression et du volume, comme une loi de compression adiabatique simple

$$P(\theta) = \frac{P_{in} \cdot V_{in}}{V(\theta)}. \quad (9.7)$$

La phase essentielle, celle de la combustion-détente est définie par deux calculs successifs.

$$Q(\theta) = Q_{inj} * f(\theta) \quad \text{et} \quad f(\theta) = 1 - \exp \left[-a \left(\frac{\theta - \theta_0}{\Delta\theta} \right)^{m+1} \right]. \quad (9.8)$$

Le premier calcul fournit la quantité de chaleur injectée $Q(\theta)$ en fonction de la quantité de chaleur globale Q_{inj} alors que la fraction brûlée $f(\theta)$ est marquée par le début de la combustion θ_0 , la durée de la combustion $\Delta\theta$, ainsi que m et a , des coefficients de vitesse de combustion. En tenant compte de l'énergie dégagée, l'expression pour la pression cylindre est

$$P(\theta) = \frac{P_{in} V_{in}}{V(\theta)} + A \cdot \frac{P_{dc} \cdot V_{dc} \cdot \frac{dQ(\theta)}{d\theta}}{V(\theta)}, \quad (9.9)$$

avec P_{dc} et V_{dc} respectivement la pression et le volume au début de la combustion. Les formules (B.9), (9.7) et (9.9) déterminent les pressions cylindres et fournissent les signaux d'excitation à l'entrée du processus. En modifiant les paramètres de la combustion, on peut modifier le son de synthèse final. L'étape suivante estime les fonctions de transfert vibratoire.

Transferts Les transferts entre la pression cylindre et les vibrations de l'enveloppe du moteur sont estimées en trois étapes : mesures, post-traitement des données et construction des filtres. En premier sont mesurées les réponses de la structure à partir du champ vibratoire, créé par la pression de la combustion à différents régimes et charges moteur. En deuxième sont détectées seulement les zones de combustion, en fournissant deux matrices : l'une avec le signal d'excitation, issu de la pression cylindre, l'autre, avec les réponses vibratoires. La combustion est considérée au-dessous de 5 kHz, afin de limiter les contributions des autres sources moteur. En dernier sont estimés des filtres FIR ou IIR, correspondant au transfert vibratoire. En notant les fonctions de transfert entre le cylindre 1, dans le point 1, avec $(\frac{\gamma}{P})_{cyl1 \rightarrow P_1}$, on obtient les vibrations du point P_1 , excité par la pression dans le cylindre 1,

$$\gamma_{cyl1 \rightarrow P_1} = P_{cyl1} * \left(\frac{\gamma}{P} \right)_{cyl1 \rightarrow P_1} \quad (9.10)$$

où $*$ désigne le produit de la convolution. Les vibrations dans le point P_1 contiennent aussi des contributions de tous les cylindres, ce qui nécessite leur addition. La formule

$$\gamma_{P_m} = \sum_n^{1 \rightarrow n_{cyl}} P_{cyl n} * \left(\frac{\gamma}{P} \right)_{cyl n \rightarrow P_m} \quad (9.11)$$

est la généralisation pour tous les n cylindres et pour tous les m points.

Rayonnement acoustique Le rayonnement acoustique peut être obtenu par deux moyens : par calcul les fonctions de transfert à partir des mesures vibro-acoustiques ou par calcul numérique, en se servant de la fonction de Green tridimensionnelle en champ libre. Pour obtenir expérimentalement le rayonnement acoustique, la contribution de tous les cylindres dans chaque point est calculée séparément et ensuite additionnée aux contributions des autres points. Les fonctions de transfert entre les vibrations et les points d'écoute sont

$$\left(\frac{P_{acoustique}}{\gamma_{P_m}} \right)_{P_m \rightarrow mic} \quad (9.12)$$

L'expression complète de la pression acoustique est donnée en fonction des pression cylindres comme

$$P_{acoustique} = \sum_n^{1 \rightarrow N_p} \left[\left(\sum_n^{1 \rightarrow n_{cyl}} P_{cyl n} * \left(\frac{\gamma}{P} \right)_{cyl n \rightarrow P_m} \right) * \left(\frac{P_{acoustique}}{\gamma_{P_m}} \right)_{P_m \rightarrow mic} \right] \quad (9.13)$$

Le deuxième moyen plus complexe pour calculer le rayonnement acoustique passe par l'utilisation de la fonction de Green. La fonction de Green permet d'exprimer la pression acoustique dans tout point extérieur M de la surface, en connaissant les accélérations vibratoires normales dans tout point de l'enveloppe vibrante du moteur M_0 . Cette approche aboutit, sous certaines simplifications et considérations physiques [158], à deux expressions de la pression acoustique dans le point M , respectivement P_b pour les basses fréquences et P_h pour les hautes fréquences :

$$P_b(M, t) \approx \sum_{i=1}^N \frac{\rho_0}{4\pi r_i} \left[\left[\frac{L}{r_i} \cos(\theta_i) + 1 \right] \gamma_n(M_i, t_n) + \frac{L}{c} \gamma_n(M_i, t_{r_i}) \cos(\theta_i) \right] \Delta S_i, \quad (9.14)$$

$$P_h(M, t) \approx \sum_{i=1}^N \frac{\rho_0}{4\pi r_i} [[\cos(\theta_i) + 1] \gamma_n(M_i, t_n)] \Delta S_i. \quad (9.15)$$

Les considérations pour les basses et les hautes fréquence sont respectivement $f \ll \frac{c}{L}$ et $f \gg \frac{c}{2\pi r_{min}}$, où r_{min} est la distance minimale entre le point d'écoute et un point M_0 de l'enveloppe vibrante du moteur. Finalement, la pression acoustique est obtenue par

$$P(M, t) = P_b(M, t)_{filtrée}[0 \rightarrow c/L] + P_h(M, t)_{filtrée}[c/2\pi r_{min} \rightarrow F_{max}] \quad (9.16)$$

où la fréquence maximale F_{max} est comprise entre 5 et 6 kHz, la limite supérieure fréquentielle où s'étend la combustion.

Avantages et limites de la méthode L'avantage principal de cette méthode est la possibilité de contrôler les paramètres de la combustion comme les caractéristiques géométriques des pistons, bielles, vilebrequin et soupapes, ainsi que les réglages d'admission, échappement et combustion du moteur. Il est possible alors de suivre les événements qui apparaissent dans le cycle moteur en fonction de l'angle vilebrequin, et de les comparer avec les bruits de synthèse correspondants. L'évaluation de la perception sonore des sons de synthèse par rapport à la mesure est faite grâce à une comparaison par paire, en termes de ressemblance perceptive [157]. En revanche, le signal de synthèse ne contient que la composante de la combustion, alors que le signal mesuré comporte également les contributions des autres sources moteur. Par conséquent, il n'est pas réaliste de s'attendre à ce que les spectres des signaux de mesure et de synthèse soient identiques. Une amélioration du modèle consiste à ajouter les autres sources vibroacoustiques comme les efforts d'inerties et les chocs internes.

9.4 Synthèse granulaire

L'idée de la synthèse granulaire est donnée par Gabor en 1946, mais le vrai début dans son utilisation, en vue d'une création musicale, est fait par Xenakis en 1971. Plus tard, Roads crée en 1978 le premier logiciel de synthèse granulaire, et en 1986 Truax rend possible l'implémentation temps-réel de la synthèse granulaire. Ces avancées ont été accompagnées par un effort de fonder la base théorique de la synthèse granulaire et de comprendre l'effet des paramètres de contrôle sur les phénomènes sonores engendrés. Dans cette section, la synthèse granulaire va être présentée d'un point de vue expérimental. La motivation principale porte sur les résultats perceptifs, notamment le masquage temps-fréquence, qui se prêtent à une analogie avec la synthèse granulaire des bruits moteur.

9.4.1 Classification des méthodes

La synthèse granulaire trouve ses origines dans le concept de "quantum de son" de Gabor, où le grain temps-fréquence représente une unité de base. Ceci exprime un esprit commun avec les méthodes d'analyse temps-fréquence et les décompositions atomiques du signal.

Deux approches de synthèse La classification des méthodes de synthèse granulaire faite par Roads distingue deux approches selon la façon d'organiser les grains [159]. Ces deux approches ont déjà été citées : synthèse avec un signal de référence et synthèse sans signal de référence. A la première catégorie de synthèse granulaire appartiennent les décompositions en atomes temps-fréquence ou en temps-échelle (ondelettes), qu'on a décrites précédemment. La méthode d'analyse-synthèse développée dans la présente thèse exploite cette facette de la synthèse granulaire.

La deuxième approche n'utilise pas de son de référence, mais dispose les grains d'une manière libre afin de créer des textures sonores expérimentales. On distingue ici deux types : la synthèse granulaire quasi-synchrone (QSGS - Quasi-Synchronous Granular Synthesis) et la synthèse granulaire asynchrone (AGS - Asynchronous Granular Synthesis). Ces deux types de synthèse utilisent comme paramètres de contrôle l'amplitude, la densité, la hauteur et le décalage en temps ("inter-onset time") entre les grains. La principale différence réside dans le fait que dans la synthèse quasi-synchrone les grains forment un ou plusieurs flux continus, alors que dans la synthèse asynchrone, les grains sont organisés en nuages asynchrones, disposés d'une manière discontinue.

Synthèse granulaire quasi-synchrone Les grains dans la synthèse quasi-synchrone se suivent, en formant un ou plusieurs flux sonores. Le décalage temporel entre les grains peut être constant ou obtenu à l'aide d'une relation linéaire en temps. Quand le décalage temporel entre les grains est régulier (constant), on obtient une fonction périodique qui permet l'interprétation de la synthèse quasi-synchrone comme un cas particulier de la modulation d'amplitude. Par ailleurs, la modulation d'amplitude apparaît lorsqu'on modifie l'amplitude d'un signal, appelé porteur avec la forme d'un autre signal, appelé modulateur. La modulation va créer une série de bandes latérales autour de la fréquence porteuse qui vont s'ajouter au spectre final. La modulation d'amplitude et le décalage temporel entre les grains combinent leurs effets en changeant la périodicité globale du son. La constatation faite par Truax révèle que si on accroit le décalage temporel entre les grains, de la valeur minimale de 0 ms jusqu'à une certaine proportion de la durée du grain, on peut entendre divers effets de modulation [160]. La répétition périodique des grains provoque une régularité qui fait apparaître dans le spectre des résonances comparables avec les régions formantiques, d'après Rocha Iturbide [161]. Le changement de régularité de répétition des grains désynchronise leurs phases et provoque des interactions exprimées par le renforcement ou l'annulation des amplitudes des grains. Au moment où le décalage temporel entre les grains devient trop important,

la structure quasi-synchrone devient irrégulière et on entend d'abord des textures stochastiques, puis des événements séparés. Ce changement marque le passage vers la synthèse asynchrone.

Synthèse granulaire asynchrone Les grains dans la synthèse asynchrone, d'après Roads, ne sont plus générés selon le concept de flots linéaires. Les grains sont distribués plutôt d'une façon statistique, en créant des nuages de grains dans une région limitée du plan temps-fréquence [159]. Les paramètres de contrôle de la synthèse asynchrone sont : le début et la durée du nuage, la durée du grain (en moyenne 5-50 ms), la densité (nombre de grains par seconde), la bande de fréquence du nuage, l'enveloppe de l'amplitude du nuage, la forme d'onde des grains et enfin, la dispersion spatiale du nuage.

9.4.2 Paramètres du grain

Le grain est l'unité de synthèse de la synthèse granulaire. Par conséquent, les paramètres du grain vont déterminer en grande partie le résultat de synthèse. Le grain est caractérisé à la fois par des informations temporelles comme le début, la durée, la forme d'enveloppe et la forme d'onde, ainsi que par des informations spectrales, comme sa fréquence. La modification de chacun de ces attributs influence les sons de synthèse et témoigne de la flexibilité de la synthèse granulaire.

Durée La durée du grain sonore est étroitement liée à sa fréquence et au nombre de grains utilisés. Plus la durée est courte, plus de grains seront nécessaires pour créer la texture sonore. La relation entre la durée et la fréquence du grain est remarquée par Gabor qui souligne la signification physiologique et perceptive de la dualité temps-fréquence de son "quantum sonore". Les expériences des années 1930 ont démontré que l'oreille traite le son comme du bruit si une oscillation sinusoïdale est écoutée pour seulement quelques cycles. En effet, pour associer une note en fréquence au grain sonore, une durée minimale de 10 ms est requise [76]. Même dans l'exemple classique de la sinusoïde, une durée minimale en temps est requise pour percevoir sa hauteur tonale.

Choix de la durée La durée du grain est fixée aussi pour des raisons de discrimination temporelle de l'oreille ainsi que de temps de calcul. Aux débuts de la musique électronique expérimentale, Moles établit une limite supérieure pour la durée de 60 ms [162]. Toutefois, Rocha Iturbide suggère l'utilisation des grains plus longs afin d'exprimer des notes entières [161]. Roads préconise une durée des grains plus efficace entre 10 et 20 ms. Cependant, Green montre que la résolution temporelle de l'oreille peut aller jusqu'à une durée des grains très petite, de l'ordre de 1-2 ms [163]. Des durées encore plus courtes (allant jusqu'à 1/1000 s) sont également possibles, mais ceci demanderait un nombre de grains plus élevé et par conséquent un calcul important. La durée moyenne du grain est comprise entre 5 et 7 ms et pour des raisons de temps de calcul, ne dépasse pas en général 8 ms dans les applications temps-réel.

Enveloppe Le processus de choix de l'enveloppe du grain est connu comme "fenêtrage". L'enveloppe du grain détermine la durée et l'amplitude du grain. L'enveloppe d'un grain est caractérisée par cinq paramètres : l'attaque, la partie soutenue et l'atténuation (connu en anglais comme "attack-sustain-decay"), ainsi que la durée totale et l'amplitude maximale de l'enveloppe (**Fig. 9.11, a**).

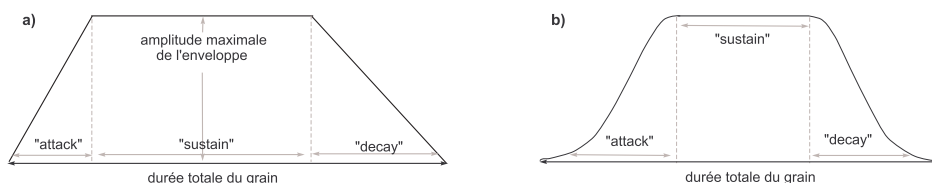


FIG. 9.11 – Schéma classique de l'enveloppe du grain (a), L'enveloppe quasi-gaussienne proposée par Roads (b)

L'attaque est une partie nécessaire de l'enveloppe du grain, car elle permet d'éviter un déclenchement sonore trop brusque. Plus l'attaque est courte, plus l'amplitude maximale du grain est atteinte rapidement et, associée à un spectre large bande, elle provoque un click au début du grain. Au contraire, un temps d'attaque large permet un début de grain en douceur. Le sustain détermine le support effectif du grain, i.e sa durée à une amplitude maximale. Si l'enveloppe possède un sommet peu étendu par rapport à la durée totale du grain, le grain est perçu d'une courte durée. L'atténuation de l'enveloppe va déterminer la vitesse avec laquelle le grain va s'arrêter. Il n'est pas obligatoire qu'elle soit symétrique à celle de

l'attaque. Les effets d'enveloppe asymétrique sont perçus d'une façon plus significative dans la répétition périodique des grains plutôt que dans un grain isolé. D'une manière générale, l'enveloppe et la durée du grain sont indépendants de la fréquence. Toutefois, ces paramètres sont liés dans le cas des ondelettes, introduites pour la synthèse granulaire par Boyer et Kronland-Martinet [164].

Choix de l'enveloppe L'enveloppe classique est la gaussienne, proposée par Gabor comme un choix optimal du grain sonore [70]. Toutefois, en raison de son sommet peu étendu, la gaussienne se révèle pas suffisamment puissante pour créer certains effets sonores. D'un autre côté, l'enveloppe rectangulaire utilisée par Xenakis possède un support effectif maximal, mais génère des discontinuités qui se traduisent par des transitoires indésirables. Roads propose un compromis entre ces deux solutions - une fenêtre quasi-gaussienne dont le début et la fin suit la courbe de la gaussienne, mais possède un sommet plus soutenu (**Fig. 9.11, b**). Cette enveloppe se rapproche en support effectif de l'enveloppe trapézoïdale. Enfin, l'enveloppe triangulaire est simple à réaliser, mais manque de partie soutenue, alors que pour des implémentations temps-réel, il est préférable d'utiliser des simples segments de lignes [165].

Forme d'onde La forme d'onde est un paramètre qui va influencer radicalement le timbre du son de synthèse. La forme d'onde peut être obtenue de différentes manières. La forme d'onde la plus simple est la forme sinusoïdale. Toutefois, le grain qui possède une forme d'onde sinusoïdale diffère d'une sinusoïde par rapport à son support fini en temps. Cet exemple montre par ailleurs le lien réciproque entre la synthèse additive et la synthèse granulaire. D'un côté, la synthèse granulaire peut être vue comme une synthèse additive, où on somme des grains à la place des sinusoïdes. D'un autre côté, la synthèse additive peut être vue comme un cas particulier de la synthèse granulaire, où les grains ont une forme d'onde sinusoïdale et un très grand support en temps. Par ailleurs, l'onde peut avoir un spectre beaucoup plus riche que la forme sinusoïdale. Elle peut être obtenue aussi à l'aide des méthodes de modulation d'amplitude (AM) ou de modulation de fréquence (FM). La forme d'onde empruntée à partir des échantillons de sons enregistrés est la technique de l'échantillonnage granulaire (Granular Sampling) [166].

Forme d'onde formantique Une forme d'onde particulière est la fonction d'onde formantique (FOF). Elle est développée dans le projet CHANT par Rodet dans le but de trouver une technique de synthèse capable de reproduire des sons complexes, notamment la voix humaine [167]. A la base de la FOF repose la modélisation de l'ensemble du conduit vocal comme un tuyau à N résonateurs, dont la fonction de transfert contient N maxima, connus comme des formants. Le formant correspond alors une forme d'onde qui possède un maximum dans une région spectrale particulière. Les FOF peuvent être obtenues en multipliant une sinusoïde par une enveloppe qui comporte un cosinus dans la partie attaque de l'enveloppe et une exponentielle décroissante pour la partie atténuation. L'avantage de la FOF est d'abord la pertinence perceptive des paramètres, la grande précision temporelle et spectrale, ainsi que la facilité d'implémentation temps réel. Rodet et d'Alessandro proposent un système d'analyse-synthèse, basé sur la modélisation par prédiction linéaire [168], alors que Clarke explore les FOF d'une manière discontinue pour la synthèse granulaire asynchrone [169]. Enfin, Eckel et Rocha Iturbide créent le logiciel GiST (Granular Synthesis Toolkit) qui permet d'effectuer la synthèse d'un point de vue spectral, mais également temporel [170].

Amplitude L'amplitude détermine l'intensité sonore du grain. De plus, l'amplitude du grain apporte un changement beaucoup plus significatif dans le contexte d'un nuage asynchrone de grains que dans le cas d'un grain isolé. Un changement d'amplitude sur les grains d'un nuage produit un effet de volume, alors qu'une amplitude constante rend le son plus applati. Roads propose un contrôle d'amplitude avec une seule enveloppe, qui gère l'amplification ou la décroissance en dB du nuage. Cependant, pour la création d'effets sonores plus intéressants, Xenakis utilise un contrôle stochastique sur l'amplitude de chaque grain. Dans ce cas, le nuage de grains est défini à l'aide d'une matrice qui contrôle indépendamment la fréquence et l'intensité de chaque grain. Ce contrôle est fait à chaque instant en temps, défini par une trame. Une série de trames va marquer l'évolution des grains dans le temps, ainsi que leurs fréquences et amplitudes.

9.4.3 Phénomènes psychoacoustiques

Les chercheurs et les compositeurs, qui ont expérimenté les différentes techniques de la synthèse granulaire, ont abouti à des résultats très intéressants sur la perception auditive. J'ai choisi de décrire ici quelques phénomènes psychoacoustiques concernant l'effet de la durée des grains, l'effet de la bande de fréquence, la densité des grains et la texture sonore, ainsi que le concept de continuum entre une perception globale du timbre et une perception locale des composantes du grain.

Effet de la durée des grains Comme il a été mentionné, les premières bornes de la durée des grains ont été définies par Gabor, Moles et Greens et comprennent une durée de 10-60 ms. Roads propose une durée moyenne de 10-20 ms, en élargissant plus tard les limites à 5-100 ms. Rocha Iturbide va encore plus loin, en proposant une durée de 1-200 ms et même au-delà. D'un point de vue psychoacoustique, Olson détermine une durée minimale de 13 ms pour percevoir la hauteur d'un son isolé. Une durée d'environ 40-50 ms est nécessaire pour que le timbre des sons isolés puisse émerger. Par ailleurs, la durée de 50 ms est mentionnée par Truax comme une valeur qui a le sens d'un seuil perceptif.

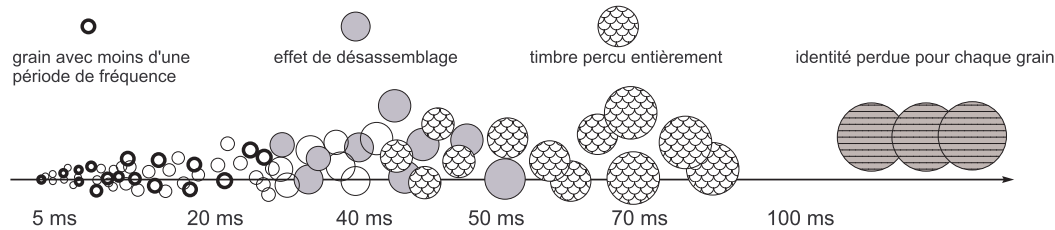


FIG. 9.12 – Notion de continuum en fonction de la durée des grains

Ce seuil marque un continuum entre la fusion des grains et leur perception comme des événements discrets (**Fig. 9.12**). Plus précisément, si la durée des grains dans un nuage est au-dessous de 50 ms, le résultat perceptif est une fusion des grains dans une texture continue. Si les grains ont une durée supérieure à 50 ms, ils sont perçus comme des événements séparés. Truax interprète ces résultats en termes de perception partielle de l'onde du grain, qui domine au-dessous de 50 ms, alors que les propriétés du timbre du grain dominant au-dessus de 50 ms. Par rapport au seuil de 50 ms, Rocha Iturbide précise qu'il ne s'agit pas d'une valeur absolue, car en effet la perception des grains va aussi dépendre de leur densité et de l'étendue de la bande de fréquence. La **Fig.9.13** illustre les valeurs perceptives de la durée des grains, comparées à la durée d'un quart de cycle moteur.

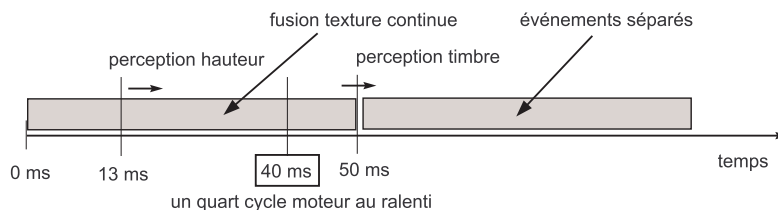


FIG. 9.13 – Quelques durées caractéristiques des grains

Effet de la bande de fréquence L'idée de continuum entre la fréquence et le bruit est expliquée par Truax de la manière suivante. Une étendue en fréquence étroite permet au système auditif de percevoir une fréquence, alors qu'une étendue au-delà de 10 % de la fréquence centrale fait disparaître la sensation de fréquence, en laissant apparaître une sensation de bruit à bande étroite. En augmentant la bande, on obtient un bruit large bande. Il existe un lien intrinsèque entre la bande de fréquence et la densité des grains. Pour une densité large, Roads établit les effets sonores en fonction de la largeur de la bande fréquentielle. Une bande étroite crée des flots de grains avec une hauteur déterminée et un spectre formantique, tout à fait comme dans la synthèse granulaire quasi-synchrone. Les bandes moyennes, comprenant un intervalle de plusieurs demi-tons, donnent des bruits colorés, alors que les bandes larges créent des nuages sonores massifs. Généralement, la relation entre la bande de fréquence et la densité est la dépendance majeure créant des effets de texture dans la synthèse granulaire.

Densité des grains La densité est définie normalement comme le nombre de grains par seconde. Cependant, dans le cas où la durée des grains change au cours du nuage asynchrone, cette mesure n'est plus valable. Roads détermine trois types de densités en fonction du rapport au silence : clairsemée, où plus de 50 % de la durée du nuage est silence ; remplie, où le nuage est rempli par des grains sonores et dense, où il existe un haut degré de recouvrement entre les grains dans le nuage. La densité est alors reliée à la durée des grains et au décalage temporel entre eux. De même, il n'est pas possible de séparer la densité de la bande de fréquence. Une étude de Mackey établit les différents seuils de perception en

fonction de la densité des grains dans un nuage. Avec 15 gr/s (grains par seconde), les structures rythmiques commencent à disparaître, alors qu'avec une succession encore plus rapide, on perd l'accélération du tempo et augmente l'intensité perçue du flot des grains. Avec 25 gr/s la perception d'ordre relatif des grains est perdue et les limites supérieure et inférieure de la bande de fréquence sont plus clairement perçues. A 50 gr/s il n'y a plus de perception des fréquences composantes internes de la texture et quand les textures deviennent de plus en plus denses, on fait l'expérience auditive d'une masse sonore. Au-delà de 100 gr/s, il apparaît un effet d'augmentation de volume et de distorsion de la masse sonore. Truax explique le phénomène de volume sonore comme la amplitude perçue du son qui tend à augmenter avec la richesse spectrale, la réverbération, la durée et l'intensité [171].

Concept de texture sonore La relation entre la densité des grains et la bande de fréquence influence grandement la texture sonore des nuages de grains. La texture sonore est un phénomène auditif qui présente un nuage de grains comme une image sonore globale stable, composée d'une multitude d'événements. MacKey compare la texture sonore à l'image visuelle d'une carpe avec un nombre large de crins et de fils [172]. Selon Smalley, la perception de la texture sonore commence au moment où l'oreille ne peut plus différencier les composantes séparées et privilégie une écoute plus globale, afin de suivre la morphologie du son [173]. Cependant, l'oreille peut toujours basculer d'une écoute globale vers une écoute de la micro-structure du son, par exemple quand de soudains changements se produisent dans la structure du son. Au moment où ces changements commencent à se répéter d'une manière stable, l'oreille passe de nouveau vers une écoute globale.

Typologie spectrale Afin d'illustrer le passage entre les deux domaines de sons, appelés spectral et morphologique, Smalley crée une typologie spectrale de passage entre le spectre et le timbre [173]. Il définit le spectre comme l'ensemble des fréquences perceptibles, en divisant le domaine de la fréquence en "ton" et "timbre". Smalley crée un schéma de perception constitué de trois références perceptives - la "note", le "nœud" et le "bruit". La note est associée attribue plus d'attention à la fréquence. La note selon Smalley est constituée d'une note propre (une fréquence bien définie), le spectre harmonique qui émerge naturellement de la note propre et enfin, le spectre inharmonique. Le nœud représente un état d'écoute intermédiaire entre la note et le bruit. Plus loin, au moment où l'auditeur ne peut plus distinguer les différents partiels, on passe d'une écoute qui sépare bien les composantes vers une écoute globale, morphologique, où les composantes fusionnent d'un point de vue perceptif. Toutefois, ce processus n'est jamais séparé du domaine temporel.

Flux des grains ("streams") Dans la création de textures sonores, les flux désignent des groupes de grains qui ont une identité en commun. Cette identité peut être par exemple l'échantillon sonore à partir duquel sont créés les grains appartenant à un flux. Le concept des flux dans la synthèse granulaire est étroitement lié à la théorie des scènes auditives. En effet, les flux auditifs marquent le passage de la synthèse quasi-synchrone vers la synthèse synchrone, créant une sorte d'état intermédiaire, désigné par continuum. Dans ce contexte, McAdams et Bregman ont travaillé sur les phénomènes de flux auditifs, formés par des événements sonores qui possèdent une grande densité, ainsi que sur la limite pour laquelle surgissent les phénomènes de séparation ou fusion entre les flux [174]. Ces aspects des scènes auditives vont être décrits dans la troisième partie de la thèse.

Perspectives Les études sur la densité d'un nuage des grains peuvent apporter une vision supplémentaire sur la notion d'indépendance et de parcimonie des atomes dans le plan temps-fréquence. Il me semble que dans cette direction on devrait chercher un critère perceptif qui serait intégré dans les algorithmes de décompositions atomiques. Par ailleurs, les études de la synthèse granulaire sur l'effet de la bande de fréquence, la durée des atomes et la formation des flux auditifs, permettent un lien intéressant avec les tests psychoacoustiques effectués sur les bruits moteur de synthèse.

Conclusion sur la synthèse sonore

Dans ce chapitre j'ai présenté des méthodes utilisées couramment pour la synthèse sonore. La technique la plus populaire, la synthèse additive, donne de très bons résultats sur les bruits moteurs. La modélisation physique fournit un calcul très précis et en relation avec la réalité physique des processus dans le cycle moteur. J'ai présenté également une autre facette de la synthèse granulaire, non pas d'un point de vue des méthodes d'analyse-synthèse, mais plutôt dans l'esprit exploratoire des compositeurs de la synthèse granulaire en vue de la création de différentes textures sonores. Par ailleurs, leurs recherches ont révélé des phénomènes psychoacoustiques intéressants, dont certains pourront être intégrés dans les méthodes de décompositions atomiques des signaux.

Orientation de la thèse : La synthèse additive semble moins adaptée pour la modélisation du bruit moteur au ralenti, car ce type de bruit a une nature davantage impulsive que spectrale. La synthèse par modèle physique exige un temps de calcul important et gère un grand nombre de paramètres. La synthèse granulaire (avec un signal de référence) apparaît donc comme un choix judicieux pour la synthèse des bruits moteur au ralenti. Elle semble un bon compromis en vue de la nature impulsive des signaux moteur, de l'interprétation des événements temporels dans le cycle moteur et la gestion des paramètres du modèle. La synthèse granulaire permet un lien intéressant avec les phénomènes psychoacoustiques.

L'analyse-synthèse du ralenti moteur

Les chapitres précédents ont exposé quatre domaines du traitement du signal : les représentations temps-fréquence, la réallocation, les décompositions atomiques et la synthèse sonore. Ce chapitre aboutit à l'application de l'ensemble de ces connaissances pour les bruits moteur au ralenti, en fournissant une nouvelle méthode d'analyse-synthèse. Cette méthode est fondée sur quatre considérations principales : le modèle granulaire (atomique) du signal détermine les paramètres à estimer ; l'expression analytique des opérateurs de réallocation à l'aide des dérivées partielles de la transformée de Gabor, donne une estimation directe des paramètres du modèle ; les critères de sélection des atomes permettent de retenir un nombre réduit d'atomes et enfin, la synthèse granulaire reconstruit le signal par la sommation des atomes. Ces points vont être discutés dans ce chapitre et seront illustrés avec des exemples de base, notamment des signaux constitués de un ou deux atomes temps-fréquence. L'application de la méthode pour les bruits moteur au ralenti conclut ce chapitre.

10.1 Bases de la méthode d'analyse-synthèse

Cette section présente le modèle atomique du signal, sa justification par rapport aux bruits moteur au ralenti, ainsi que le schéma de l'algorithme d'analyse-synthèse développé dans la thèse.

10.1.1 Modèle atomique du signal

L'analyse temps-fréquence a montré que, du fait de leur caractère répétitif, les bruits moteur possèdent une structure à la fois impulsive et harmonique. On a également discuté la pertinence de la synthèse additive qui s'adapte naturellement à la structure harmonique des bruits moteur aux régimes moyens et élevés. Cependant, le ralenti moteur révèle une nature davantage impulsive, qui semble ne pas être facilement modélisable par la synthèse additive. Dans l'objectif de synthétiser les sons impulsifs répétitifs, dont le ralenti moteur fait partie, on adopte un modèle granulaire.

Modèle Le modèle granulaire (appelé aussi modèle atomique) considère le bruit moteur comme une somme finie d'atomes temps-fréquence. Le bruit de moteur s_{moteur} est modélisé à l'aide de N atomes de type chirp linéaire à enveloppe gaussienne. Le modèle se résume avec l'expression

$$s_{moteur}(t) = \sum_{k=1}^N A_k e^{-\frac{(t-t_k)^2}{2\sigma_k^2}} e^{i[\omega_k(t-t_k) + \frac{\beta_k}{2}(t-t_k)^2]}. \quad (10.1)$$

Les atomes chirp ont l'avantage de s'adapter mieux que les atomes non chirpés à la structure locale du signal. La fréquence instantanée d'un atome chirp satisfait la loi linéaire $\omega + \beta t$, où le paramètre de chirp β traduit la pente (l'orientation) de l'atome dans le plan temps-fréquence. L'enveloppe gaussienne est préférée devant d'autres fonctions (Hann, Hamming, etc.) car, outre sa propriété de localisation temps-fréquence, elle permettra le calcul analytique des paramètres du modèle (Eq. 10.1).

Paramètres Cinq paramètres identifient l'atome chirp (Fig. 10.1) :

- t_k : le sommet temporel de l'enveloppe de l'atome,
- ω_k : la fréquence instantanée de l'atome (à $t = 0$),
- σ_k : le paramètre de support temporel de l'enveloppe gaussienne,
- β_k : le paramètre de chirp (orientation),
- A_k : l'amplitude de l'atome.

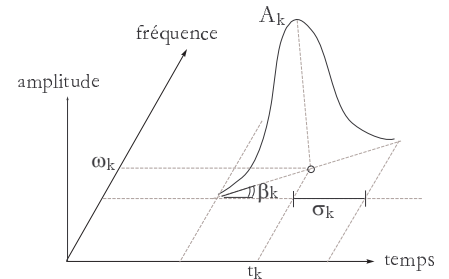


Fig. 10.1 – Illustration des paramètres d'un atome chirp temps-fréquence

Objectif L'objectif de l'algorithme d'analyse-synthèse sera d'analyser un seul cycle moteur, porteur de l'information essentielle du bruit moteur et de le reconstruire à l'aide d'un nombre limité d'atomes

chirp. Le processus d'analyse-synthèse comprend l'estimation des paramètres des atomes selon le modèle granulaire (Eq. 10.1). Le bruit moteur final sera reconstruit par périodisation des atomes constituant un cycle moteur de resynthèse.

Atomes chirps et bruit moteur La transformée de Gabor, décrite précédemment, est un outil d'analyse qui décompose les signaux sonores sur une famille de fonctions temps-fréquence. Dans la définition originale de Gabor, ces fonctions sont des atomes gaussiens non chirpés, représentant le signal à l'aide d'un pavage "vertical" ou "horizontal" du plan temps-fréquence. Toutefois, la structure de nombreux signaux, comme par exemple les cris d'oiseaux, appelés "chirp" ("gazouillis" en français), comportent une structure plutôt "diagonale" dans le plan temps-fréquence. Cette structure traduit une loi d'évolution de la fréquence instantanée.

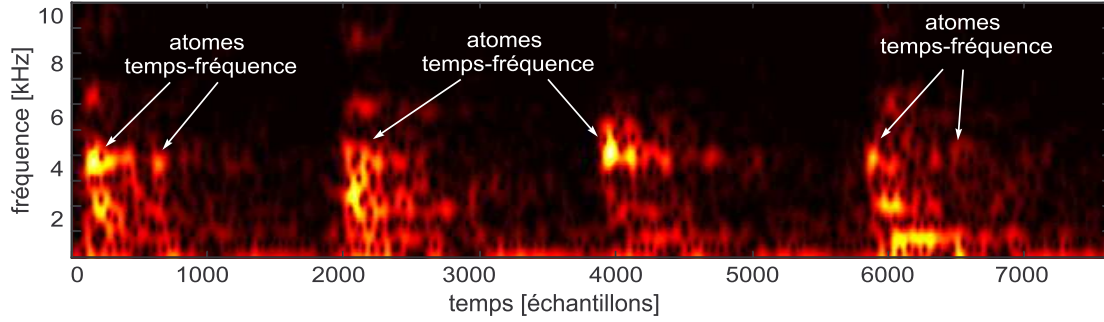


FIG. 10.2 – La structure atomique d'un cycle moteur au ralenti

Des structures similaires du type "atomes chirp" peuvent être détectées dans l'analyse d'un cycle moteur au ralenti (Fig. 10.2). Afin de mieux décrire la structure atomique du bruit moteur au ralenti, il semble judicieux d'utiliser un modèle de signal avec des atomes chirpés, pouvant adapter leur orientation dans le plan temps-fréquence. L'objectif de l'analyse-synthèse consistera à couvrir au mieux l'image temps-fréquence d'un cycle de bruit moteur avec des atomes temps-fréquence chirpés.

Opérations dans le plan temps-fréquence Les atomes chirp représentent une extension de la famille d'atomes temps-fréquence. Comme on l'a déjà vu, chaque atome de cette famille est obtenu à partir d'un atome à enveloppe gaussienne à l'aide des opérations de translation en temps, modulation en fréquence et rotation dans le plan temps-fréquence (Fig. 10.3), et de la modification du paramètre de la gaussienne permettant un élargissement ou rétrécissement du support temporel (Fig. 10.4).

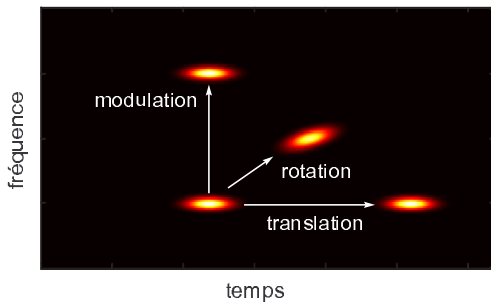


Fig. 10.3 – Les opérations de modulation, rotation et translation des atomes dans le plan temps-fréquence

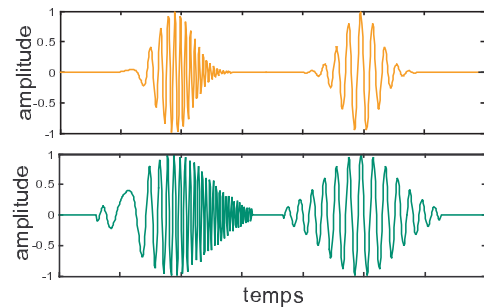


Fig. 10.4 – Partie réelle d'un atome chirp et d'un atome non chirpé avec deux supports différents

En effet, ces opérations sont associées aux 4 paramètres du modèle. Le cinquième paramètre, l'amplitude A_k , déterminera le niveau énergétique (l'intensité) de l'atome. On verra par la suite, que l'estimation de ces paramètres dépendra du choix du support de la fenêtre d'analyse.

10.1.2 Schéma d'analyse-synthèse

La présentation de la méthode dans les sections suivantes va suivre la logique de l'algorithme d'analyse-synthèse. Cet algorithme comprend plusieurs étapes, comme l'illustre le schéma (Fig. 10.5).

Étapes de l'algorithme La première étape de l'algorithme effectue l'extraction d'un cycle moteur à partir d'un bruit enregistré. Comme on l'a déjà mentionné, un seul cycle moteur va être notre unité d'analyse à partir de laquelle on peut reconstruire, par périodisation, le bruit moteur total. Ensuite, la partie essentielle de l'analyse comporte, successivement, le calcul de la transformée de Gabor et de plusieurs de ses dérivées. A partir de ces dérivées seront définis de nouveaux opérateurs de réallocation. Ils permettront la production de trois "images" temps-fréquence complémentaires, représentant respectivement l'énergie, le support et l'orientation locaux du signal. Ces représentations fourniront une première estimation des paramètres σ_k , β_k , t_k et ω_k d'un sur-ensemble d'atomes.

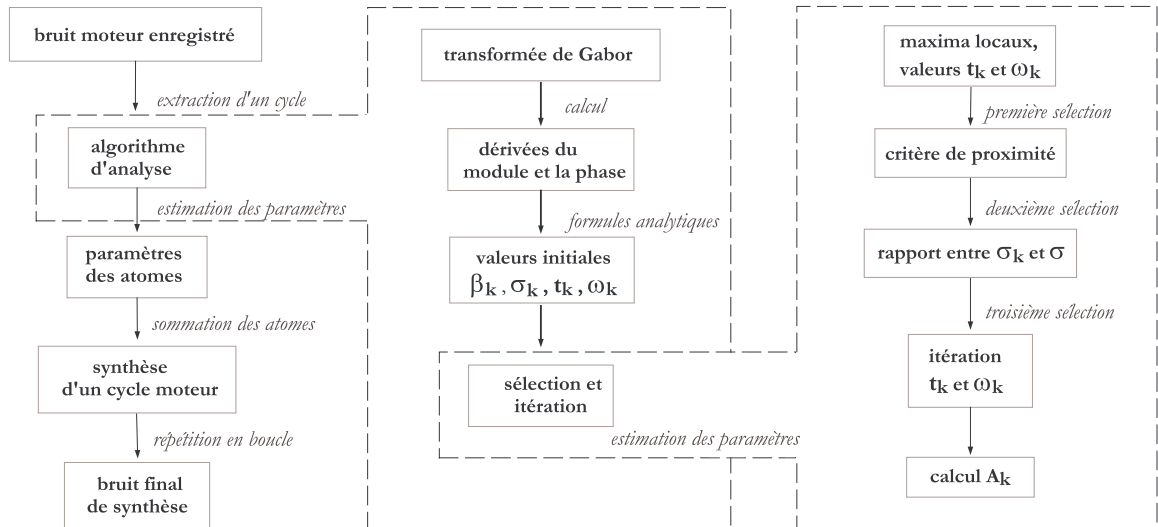


FIG. 10.5 – Le schéma de la méthode d'analyse-synthèse

Pour faire le tri des atomes, on introduit trois différents critères de sélection. Les coordonnées temps-fréquence des atomes finalement retenus sont affinées à l'aide d'un algorithme itératif. En dernier est calculée l'amplitude A_k des atomes. En disposant de tous les paramètres, on somme les parties réelles des atomes afin d'obtenir la resynthèse d'un cycle. Ensuite, en périodisant chaque atome, on reconstruit plusieurs cycles moteur.

Particularités de la méthode L'originalité de l'approche réside dans une nouvelle définition des opérateurs de réallocation qui, comme on le verra plus loin, localisent l'information du signal non plus le long de trajectoires, mais tendent à la concentrer sur des points dans le plan temps-fréquence (Fig.10.6).

Par ailleurs, l'expression analytique des opérateurs de réallocation est obtenue en utilisant une enveloppe gaussienne pour les atomes. Un point important de la méthode est la concentration encore plus forte grâce à l'élimination d'un nombre d'atomes qui ne satisfont pas certains critères de sélection. Le lien avec les méthodes d'"arête et squelette" s'effectue en deux points. Le premier point est dans la sélection initiale des atomes, qui peut se faire à partir d'un ensemble de maxima locaux de la transformée de Gabor. C'est une approche commune également à la méthode "ridge pursuit". Le deuxième point commun avec les méthodes "arête et squelette" réside dans le processus d'itération finale qui affine les valeurs de t_k et ω_k pour calculer les points fixes des opérateurs de la réallocation.

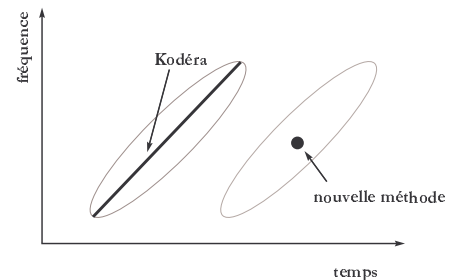


Fig. 10.6 – Différence conceptuelle entre la réallocation classique et la nouvelle méthode

10.2 Estimation des paramètres

L'estimation des paramètres des atomes représente le cœur de la méthode d'analyse-synthèse. Ici on va exploiter les avantages que nous donne le fait d'avoir des expressions explicites de la transformée de Gabor et de ses dérivées. En effet, dans le cas d'un atome gaussien, un calcul analytique nous permet d'aboutir à une estimation directe pour les 5 paramètres d'un atome chirp, σ_1 , β_1 , t_1 , ω_1 et A_1 .

10.2.1 Transformée de Gabor d'un atome non chirpé

Avant de présenter les formules analytiques pour un atome gaussien chirpé, on va considérer le cas plus simple, celui d'un atome non chirpé.

Expression analytique Le produit scalaire de l'atome à enveloppe gaussienne

$$s(t) = e^{-\frac{(t-t_1)^2}{2\sigma_1^2}} \cdot e^{i\omega_1(t-t_1)} \quad \text{avec des fonctions de Gabor} \quad h_{\tau,\alpha}(t) = e^{-\frac{(t-\tau)^2}{2\sigma^2}} \cdot e^{i\alpha(t-\tau)} \quad (10.2)$$

donne la transformée de Gabor de l'atome

$$G(\tau, \alpha) = \langle s(t), h_{\tau,\alpha} \rangle = \frac{\sqrt{2\pi}\sigma_1\sigma}{\sqrt{\sigma_1^2 + \sigma^2}} \cdot e^{-\frac{1}{2(\sigma^2 + \sigma_1^2)}(\tau-t_1)^2} \cdot e^{-\frac{\sigma_1^2\sigma^2}{2(\sigma^2 + \sigma_1^2)}(\alpha-\omega_1)^2} \cdot e^{i\left(\frac{\omega_1\sigma_1^2 + \alpha\sigma^2}{\sigma_1^2 + \sigma^2}\right)(\tau-t_1)} \quad (10.3)$$

Cette formule est obtenue grâce au fait que la transformée de Gabor d'un atome gaussien représente une intégrale gaussienne,

$$\int e^{-px^2 + 2qx} dx = \frac{\sqrt{\pi}}{p} e^{\frac{q^2}{p}}, \quad (10.4)$$

dont l'expression analytique est connue.

Lignes de module L'expression (10.3) donne une information sur les lignes de module constant qui forment une ellipse dans le plan temps-fréquence avec des côtés en temps et en fréquence respectivement

$$\frac{1}{\sqrt{2(\sigma_1^2 + \sigma^2)}} \quad \text{et} \quad \frac{\sigma_1\sigma}{\sqrt{2(\sigma_1^2 + \sigma^2)}}. \quad (10.5)$$

Dans le cas où l'atome analysé de Gabor a le même paramètre σ que la famille d'analyse ($\sigma_1 = \sigma$), on obtient la formule du noyau reproduisant

$$P_h(\tau, \tau_0, \alpha, \alpha_0) = \sigma\sqrt{\pi} \cdot e^{-\frac{1}{4\sigma^2}(\tau-\tau_0)^2} \cdot e^{-\frac{\sigma^2}{4}(\alpha-\alpha_0)^2} \cdot e^{i\left(\frac{\alpha+\alpha_0}{2}\right)(\tau-\tau_0)}, \quad (10.6)$$

qui détermine les lignes de module de la transformée de Gabor d'un atome gaussien comme une ellipse dans le plan temps-fréquence.

Fréquence locale et retard local La forme complexe de la transformée de Gabor en termes de module M et phase Φ s'écrit $G(\tau, \alpha) = M(\tau, \alpha)e^{i\Phi(\tau, \alpha)}$. La phase de la transformée de Gabor d'un atome gaussien non chirpé est exprimée comme

$$\Phi(\tau, \alpha) = \left(\frac{\omega_1\sigma_1^2 + \alpha\sigma^2}{\sigma_1^2 + \sigma^2} \right) (\tau - t_1). \quad (10.7)$$

Les dérivées partielles de la phase sont respectivement la fréquence locale et le retard local

$$\frac{\partial \Phi}{\partial \tau} = \frac{\alpha\sigma^2 + \omega_1\sigma_1^2}{\sigma_1^2 + \sigma^2} \quad \text{et} \quad \frac{\partial \Phi}{\partial \alpha} = \frac{\sigma^2(\tau - t_1)}{\sigma_1^2 + \sigma^2}. \quad (10.8)$$

Par ailleurs, il découle des équations (10.8) que pour les points $\tau = t_1$ et $\alpha = \omega_1$, on a les relations

$$\frac{\partial \Phi(t_1, \alpha)}{\partial \alpha} = 0 \quad \text{et} \quad \frac{\partial \Phi(\tau, \omega_1)}{\partial \tau} = \omega_1. \quad (10.9)$$

La dérivation de la fréquence locale donne une expression directe pour le paramètre σ_1 de l'atome,

$$\sigma_1 = \sigma \sqrt{\frac{1}{\Phi_{\tau,\alpha}} - 1}. \quad (10.10)$$

Ici on a introduit la notation de la dérivée mixte $\Phi_{\tau,\alpha} = \frac{\partial^2 \Phi}{\partial \tau \partial \alpha}$. L'estimation de σ_1 donnée par la formule (10.10) sera utilisée par l'algorithme d'estimation dans le cas où l'atome n'est pas chirpé. Par ailleurs, la dérivation de la phase (et du module) fait tomber la dépendance des paramètres de translation τ et modulation α et nous donne ainsi la possibilité d'avoir une expression directe des paramètres de l'atome.

10.2.2 Transformée de Gabor d'un atome chirp

Les formules des dérivées partielles du module et la phase de la transformée de Gabor d'un atome non chirpé montrent qu'il est possible de trouver une expression pour les paramètres de l'atome. Les dérivées partielles sont calculables numériquement à partir de la transformée de Gabor. Notre objectif maintenant va être de trouver des expressions directes pour les paramètres σ_1 et β_1 d'un atome chirpé.

Expression analytique L'atome chirpé à enveloppe gaussienne s'exprime comme

$$s(t) = e^{-\frac{(t-t_1)^2}{2\sigma_1^2}} e^{i[\omega_1(t-t_1) + \frac{\beta_1}{2}(t-t_1)^2]}. \quad (10.11)$$

L'expression analytique de la transformée de Gabor de l'atome chirp s'écrit sous la forme

$$G(\tau, \alpha) = \frac{\sqrt{2\pi}\sigma_1\sigma}{\sqrt{\sigma_1^2 + \sigma^2 - i\beta_1\sigma_1^2\sigma^2}} \cdot e^{\frac{1}{2} \frac{A + iB}{(\sigma^2 + \sigma_1^2)^2 + \beta_1^2\sigma_1^4\sigma^4}}, \quad (10.12)$$

où

$$\begin{aligned} A &= -[\sigma_1^2 + \sigma^2 - \beta_1^2\sigma_1^2](\tau - t_1)^2 + 2\beta_1(\tau - t_1)[(\sigma_1^2 + \sigma^2)2\alpha\sigma_1^2\sigma^2 - (\sigma_1^2\omega_1 + \sigma^2\alpha)] - \sigma_1^2\sigma^2(\sigma_1^2 + \sigma^2)(\alpha - \omega_1)^2, \\ B &= \beta_1[\sigma_1^2(\sigma_1^2 + \sigma^2) - 1](\tau - t_1)^2 + 2(\tau - t_1)[(\sigma_1^2 + \sigma^2)(\sigma_1^2\omega_1^2 + \sigma^2\alpha) + \alpha\beta_1^2\sigma_1^2\sigma^2] - \beta_1\sigma_1^2\sigma^2(\alpha - \omega_1)^2. \end{aligned}$$

Naturellement, si on pose $\beta = 0$ dans la formule (10.12), on obtient la formule de la transformée de Gabor d'un atome non chirpé (10.3).

Dérivées de la phase et du module On a mentionné que les dérivées de la phase de la transformée de Gabor par rapport aux paramètres de translation en temps et en fréquence, sont calculables numériquement. En effet, elles sont déduites directement à partir de la transformée à l'aide des formules

$$\frac{\partial \Phi}{\partial \tau}(\tau, \alpha) = \text{Im} \left\{ \frac{\frac{\partial G(\tau, \alpha)}{\partial \tau}}{G(\tau, \alpha)} \right\}, \quad \frac{\partial \Phi}{\partial \alpha}(\tau, \alpha) = \text{Im} \left\{ \frac{\frac{\partial G(\tau, \alpha)}{\partial \alpha}}{G(\tau, \alpha)} \right\}. \quad (10.13)$$

Ces formules sont obtenues directement à partir de l'expression de la transformée de Gabor en termes de module et phase $G(\tau, \alpha) = M(\tau, \alpha)e^{i\Phi(\tau, \alpha)}$. Les dérivées du module sont calculées d'une manière analogue, comme la partie réelle

$$\frac{\frac{\partial M}{\partial \tau}(\tau, \alpha)}{M(\tau, \alpha)} = \text{Re} \left\{ \frac{\frac{\partial G(\tau, \alpha)}{\partial \tau}}{G(\tau, \alpha)} \right\} = \frac{\partial}{\partial \tau} \ln(M), \quad \frac{\frac{\partial M}{\partial \alpha}(\tau, \alpha)}{M(\tau, \alpha)} = \text{Re} \left\{ \frac{\frac{\partial G(\tau, \alpha)}{\partial \alpha}}{G(\tau, \alpha)} \right\} = \frac{\partial}{\partial \alpha} \ln(M). \quad (10.14)$$

En pratique, dans le cas d'une fenêtre d'analyse gaussienne, il est plus pratique d'utiliser les formules suivantes pour le calcul des dérivées de la phase et du module

$$\frac{\partial \Phi}{\partial \tau}(\tau, \alpha) = \alpha + \text{Im} \left\{ \frac{G_{dh}(\tau, \alpha)}{G(\tau, \alpha)} \right\}, \quad \frac{\partial M}{\partial \tau}(\tau, \alpha) = \text{Re} \left\{ \frac{G_{dh}(\tau, \alpha)}{G(\tau, \alpha)} \right\}, \quad (10.15)$$

où $G_{dh}(\tau, \alpha)$ est la transformée de Gabor avec la fenêtre dérivée $\frac{dh(t)}{dt}$.

Le calcul à partir de la transformée de Gabor, donne les formules suivantes pour les dérivées de phase dans le cas d'un seul atome gaussien chirpé

$$\frac{\partial \Phi}{\partial \tau}(\tau, \alpha) = \frac{[\beta_1^2 \sigma^4 \alpha + (\tau - t_1) + \omega_1] \sigma_1^4 + (\omega_1 + \alpha) \sigma^2 \sigma_1^2 + \alpha \sigma^4}{(\sigma^2 + \sigma_1^2)^2 + \beta_1^2 \sigma_1^4 \sigma^4}, \quad (10.16)$$

$$\frac{\partial \Phi}{\partial \alpha}(\tau, \alpha) = \frac{[(\tau - t_1) \beta_1^2 + (\omega_1 - \alpha) \beta_1] \sigma_1^4 \sigma^4 + (\tau - t_1) \sigma_1^2 \sigma^2 + (\tau - t_1) \sigma^4}{(\sigma^2 + \sigma_1^2)^2 + \beta_1^2 \sigma_1^4 \sigma^4}. \quad (10.17)$$

Au point (t_1, ω_1) , on obtient des expressions connues pour la fréquence locale et le retard local

$$\frac{\partial \Phi}{\partial \tau}(\tau, \alpha) = \alpha \quad \text{et} \quad \frac{\partial \Phi}{\partial \alpha}(\tau, \alpha) = 0. \quad (10.18)$$

Dans la voie α , qui correspond à la fréquence du chirp, i.e. $\alpha = \omega_1$, la fréquence locale a l'expression

$$\frac{\partial \Phi}{\partial \tau}(\tau, \alpha = \omega_1) = \alpha + \frac{(\tau - t_1) \beta \sigma_1^4}{\sigma_1^2 + \sigma^2 + \beta^2 \sigma_1^2 \sigma^2}. \quad (10.19)$$

Les dérivées du module par rapport aux paramètres de translation et de modulation sont

$$\frac{\partial \ln(M)}{\partial \alpha}(\tau, \alpha) = \frac{[(\sigma_1^2 + \sigma^2)(\omega_1 - \alpha) + \sigma_1^2 \beta_1 (\tau - t_1)] \sigma_1^2 \sigma^2}{(\sigma^2 + \sigma_1^2)^2 + \beta_1^2 \sigma_1^4 \sigma^4}, \quad (10.20)$$

$$\frac{\partial \ln(M)}{\partial \tau}(\tau, \alpha) = \frac{(\sigma_1^2 + \sigma^2 + \beta_1^2 \sigma_1^4 \sigma^2)(t_1 - \tau) + (\sigma_1^4 \sigma^2 \beta_1)(\alpha - \omega_1)}{(\sigma^2 + \sigma_1^2)^2 + \beta_1^2 \sigma_1^4 \sigma^4}. \quad (10.21)$$

On peut apercevoir que les deux dérivées comportent, au premier degré, les paramètres de la transformée, α et τ . Par conséquent, une future dérivation mixte des expressions 10.20 et 10.21 éliminera la dépendance de ces deux paramètres, en laissant seulement la participation des paramètres de l'atome.

10.2.3 Estimateurs pour β_k et σ_k

L'objectif de cette section est de présenter les estimateurs pour les paramètres β_1 et σ_1 de l'atome, calculés en fonction des quantités connues et calculables du signal. Ces quantités sont la transformée de Gabor $G(\tau, \alpha)$ et ses dérivées du module et de la phase $\frac{\partial \ln(M)}{\partial \alpha}(\tau, \alpha)$, $\frac{\partial \ln(M)}{\partial \tau}(\tau, \alpha)$, $\frac{\partial \Phi}{\partial \alpha}(\tau, \alpha)$ et $\frac{\partial \Phi}{\partial \tau}(\tau, \alpha)$. Pour ce faire, on va utiliser les dérivées secondes mixtes du module et de la phase.

Les dérivées mixtes, qu'on note pour simplicité $\Phi_{\tau, \alpha} = \frac{\partial^2 \Phi}{\partial \tau \partial \alpha}$ et $M_{\tau, \alpha} = \frac{\partial^2 \ln(M)}{\partial \tau \partial \alpha}$, forment un système de deux équations

$$\Phi_{\tau, \alpha} = \frac{\sigma^4 + \sigma_1^2 \sigma^2 + \beta_1^2 \sigma^4 \sigma_1^4}{(\sigma_1^2 + \sigma^2)^2 + \beta_1^2 \sigma_1^4 \sigma^4}, \quad (10.22)$$

$$M_{\tau, \alpha} = \frac{\beta_1 \sigma_1^4 \sigma^2}{(\sigma_1^2 + \sigma^2)^2 + \beta_1^2 \sigma_1^4 \sigma^4}. \quad (10.23)$$

Le système d'équations (10.22) et (10.23) ne peut pas être résolu d'une façon simple afin obtenir directement des expressions pour σ_1 et β_1 . En revanche, on peut le résoudre par rapport aux variables σ_1^2 et β_1^2 , dont les expressions sont les suivantes

$$\sigma_1^2 = \frac{\sigma^2 M_{\tau, \alpha}}{\sigma^2 \beta_1 (1 - \Phi_{\tau, \alpha} - M_{\tau, \alpha})}, \quad (10.24)$$

$$\beta_1^2 = \frac{\beta_1}{M_{\tau, \alpha}} (\sigma_1^2 - 2\Phi_{\tau, \alpha}) - \frac{\sigma_1^4 + \sigma^4}{\sigma_1^4 \sigma^4}. \quad (10.25)$$

En remplaçant l'expression de σ_1^2 de l'équation (10.24) dans (10.25), on obtient une équation quadratique par rapport à β . Une des racines donne l'expression explicite pour β_1 . Ensuite, pour obtenir σ_1 , on

remplace dans (10.24) la valeur obtenue de β_1 , en prenant la racine carrée.

Finalement, deux formules d'estimation sont établies

$$\beta_1 = \frac{M_{\tau,\alpha}}{\sigma^2[M_{\tau,\alpha} + (1 - \Phi_{\tau,\alpha})^2]}, \quad (10.26)$$

$$\sigma_1 = \sigma \sqrt{\frac{(1 - \Phi_{\tau,\alpha})^2 + M_{\tau,\alpha}}{\Phi_{\tau,\alpha}(1 - \Phi_{\tau,\alpha}) - M_{\tau,\alpha}}}. \quad (10.27)$$

Les formules (10.26) et (10.27) sont valables et calculables en tout point (τ, α) du plan temps-fréquence. Elles peuvent donc être interprétées comme deux représentations temps-fréquence complémentaires de la transformée de Gabor $G(\tau, \alpha)$ et fournissent des images temps-fréquence du support et de l'orientation temps-fréquence locaux des différentes contributions du signal. Dans le cas d'un signal constitué d'un seul atome, elles fournissent exactement le support et l'orientation de l'atome, quelque soit le point (τ, α) du plan temps-fréquence où elles sont calculées. Cependant, il est judicieux d'évaluer ces expressions au plus près de la position (t_1, ω_1) de l'atome.

10.2.4 Opérateurs de réallocation temps-fréquence

Parallèlement au calcul des paramètres de support et d'orientation, il est possible d'obtenir les expressions analytiques des positions en temps et fréquence, t_1 et ω_1 , de l'atome. Ces expressions analytiques vont définir les nouveaux opérateurs de réallocation. De plus, elles permettront une recherche itérative afin de situer la position (t_k, ω_k) plus finement que sur les points de discrétisation en temps et en fréquence.

Pour obtenir des opérateurs de réallocation, on va se servir des expressions analytiques de la fréquence locale et du retard local, obtenues plus haut

$$\frac{\partial \Phi}{\partial \tau}(\tau, \alpha) = \frac{[\beta_1^2 \sigma^4 \alpha + (\tau - t_1) + \omega_1] \sigma_1^4 + (\omega_1 + \alpha) \sigma^2 \sigma_1^2 + \alpha \sigma^4}{(\sigma^2 + \sigma_1^2)^2 + \beta_1^2 \sigma_1^4 \sigma^4}, \quad (10.28)$$

$$\frac{\partial \Phi}{\partial \alpha}(\tau, \alpha) = \frac{[(\tau - t_1) \beta_1^2 + (\omega_1 - \alpha) \beta_1] \sigma_1^4 \sigma^4 + (\tau - t_1) \sigma_1^2 \sigma^2 + (\tau - t_1) \sigma^4}{(\sigma^2 + \sigma_1^2)^2 + \beta_1^2 \sigma_1^4 \sigma^4}. \quad (10.29)$$

En résolvant le système formé par les deux équations (10.28) et (10.29), par rapport aux paramètres t_1 et ω_1 , on aboutit aux deux formules

$$t_1 = \tau - \frac{\partial \Phi}{\partial \alpha} \left(1 + \frac{\sigma_1^2}{\sigma^2} \right) + \beta_1 \sigma_1^2 \sigma^2 \left(\frac{\partial \Phi}{\partial \tau} - \alpha \right), \quad (10.30)$$

$$\omega_1 = \frac{\partial \Phi}{\partial \tau} + \left(\frac{\partial \Phi}{\partial \tau} - \alpha \right) \left(\beta_1^2 \sigma_1^2 \sigma^2 + \frac{\sigma^2}{\sigma_1^2} \right) - \beta_1 \frac{\sigma_1^2}{\sigma^2} \frac{\partial \Phi}{\partial \alpha}. \quad (10.31)$$

Les expressions (10.30) et (10.31) décrivent les nouveaux opérateurs de réallocation dans le cas d'un atome chirp. Les paramètres β_1 et σ_1 qui participent dans le calcul de (10.30) et (10.31), sont connus d'après les formules (10.26) et (10.27). Par ailleurs, dans le cas où l'atome n'est pas chirpé, i.e $\beta = 0$, les opérateurs de réallocation se réduisent à

$$t_1 = \tau - \frac{\partial \Phi}{\partial \alpha} \left(1 + \frac{\sigma_1^2}{\sigma^2} \right), \quad (10.32)$$

$$\omega_1 = \frac{\partial \Phi}{\partial \tau} + \left(\frac{\partial \Phi}{\partial \tau} - \alpha \right) \frac{\sigma^2}{\sigma_1^2}. \quad (10.33)$$

Les formules 10.30 et 10.31, ainsi que 10.32 et 10.33, fournissent la position exacte de l'atome chirp quelque soit le point temps-fréquence où elles sont calculées. Ces nouveaux opérateurs de réallocation vont donc tendre à concentrer l'énergie du signal dans le plan temps-fréquence non plus sur des trajectoires mais sur des points. On peut remarquer que les points fixes de ces nouveaux opérateurs de réallocation $(\tau = t_1, \alpha = \omega_1)$ sont les points du plan temps-fréquence pour lesquels on a simultanément

$$\frac{\partial \Phi}{\partial \tau} = \alpha \quad \text{et} \quad \frac{\partial \Phi}{\partial \alpha} = 0. \quad (10.34)$$

Ainsi, parallèlement au spectrogramme réalloué, dans le cas d'un signal constitué d'un seul atome gaussien, cette méthode va permettre de produire trois représentations temps-fréquence. La première représentation est de type "énergétique" et consiste à utiliser les opérateurs définis par les équations 10.30 et 10.31 en lieu et place des opérateurs "classiques" de réallocation pour déplacer l'énergie du spectrogramme. Les deux autres images sont constituées des valeurs des supports (Eq.10.24) et l'orientation (Eq. 10.25) des atomes. Ces deux images complémentaires permettent d'accéder non seulement aux positions réallouées au sens du barycentre de temps et en fréquence, comme le fait la réallocation classique, mais donnent une information sur les supports et les orientations des atomes.

10.2.5 Exemples de représentation

La nouvelle méthode d'analyse-synthèse est fondée sur deux approches complémentaires, quoi que bien distinctes : qualitative et quantitative. L'approche qualitative est associée aux moyens de représentation du signal dans le plan temps-fréquence, alors que l'approche quantitative est liée à l'estimation des valeurs concrètes des paramètres. Cette section présente des exemples illustrant l'approche qualitative, en termes d'images temps-fréquence, associées aux nouveaux opérateurs de réallocation et à la répartition des valeurs du support σ_k et l'orientation β_k des atomes.

Exemple 1 : Représentation d'un atome chirp

Le signal constitué d'un seul atome chirp correspond parfaitement au modèle atomique choisi (Eq. 10.1). Les coordonnées de l'atome dans cet exemple sont $t_1 = 16,28ms$ en temps et $\omega_1 = 11025Hz$ en fréquence. Les trois images temps-fréquence de la **Fig. 10.7** permettent de comparer respectivement les représentations qualitatives : du spectrogramme de Gabor (**Fig. 10.7 a**), des opérateurs de réallocation de Kodera (**Fig. 10.7 b**) et enfin, de notre méthode (**Fig. 10.7 c**). Comme on l'a déjà vu dans le chapitre consacré à la réallocation, la concentration énergétique d'un atome chirp par les opérateurs classiques de réallocation de Kodera se fait le long d'une ligne (**Fig. 10.7 b**). Les opérateurs de réallocation de notre méthode (qu'on a définis selon les formules 10.30 et 10.31) concentrent davantage l'information de l'atome non plus sur une ligne, mais en un seul point, correspondant exactement aux coordonnées de l'atome (**Fig. 10.7 c**). Cette image qualitative va être complétée par deux autres images temps-fréquence, du support et de l'orientation de l'atome, introduites à l'aide des deux exemples suivants.

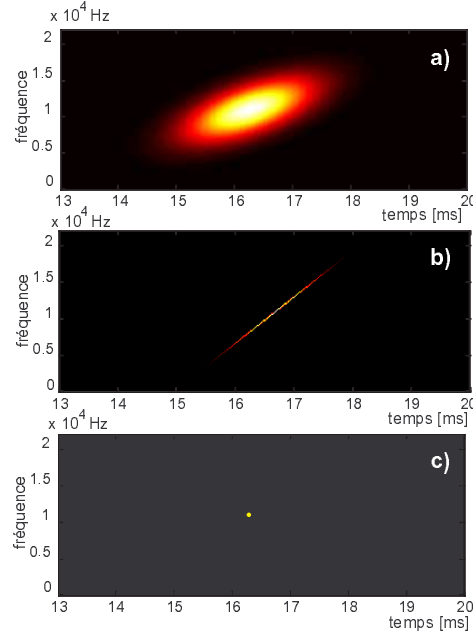


Fig. 10.7 – Représentation d'un atome chirp (détails : voir le texte)

Exemple 2 : Cartographie du support d'un atome chirp

Les formules analytiques (10.30) et (10.31) permettent de calculer les coordonnées de l'atome pour chaque point du plan temps-fréquence. De même, les formules (10.26) et (10.27) déterminent les valeurs des paramètres σ_1 et β_1 . Il est possible de tracer une image temps-fréquence du paramètre σ_1 (et de même pour le paramètre β_1), qu'on va appeler cartographie du support de l'atome. Dans le cas d'un signal composé de plusieurs atomes, cette cartographie peut donner une indication de la séparation des composantes du signal dans le plan temps-fréquence, ainsi que de l'influence de la fenêtre d'analyse sur l'estimation. Cette cartographie constitue une mesure de la "largeur de bande locale" de l'atome.

La cartographie du support d'un atome chirp permet de visualiser le support de l'atome dans le plan temps-fréquence (**Fig. 10.8**). L'atome chirp ici possède des paramètres $\sigma_1 = 20$ et $\beta = 0.04$, et il est analysé avec trois fenêtres d'analyse différentes, avec respectivement $\sigma = 12.5$, $\sigma = 25$ et $\sigma = 50$. L'étendue de la région où le support de l'atome est défini, est proportionnelle à l'étendue déterminée par le support de la fenêtre d'analyse, car les formules analytiques expriment σ_1 en fonction de σ . Toutefois, quelque soit la fenêtre d'analyse, la vraie valeur de σ_1 de l'atome est estimée correctement.

Sur toute la région, marquée par le support de la fenêtre d'analyse et celui de l'atome, la cartographie du support σ_1 est quasiment constante (**Fig. 10.8 d, e, f**). Cette constatation est logique, puisque les formules (10.26) et (10.27) sont vraies dans tout le plan temps-fréquence. On peut observer que la valeur de σ_1 est légèrement altérée au bord de sa région de définition. Cet effet provient de la troncature de la

fenêtre d'analyse. Dans le cas d'un seul atome, ceci n'est pas gênant.

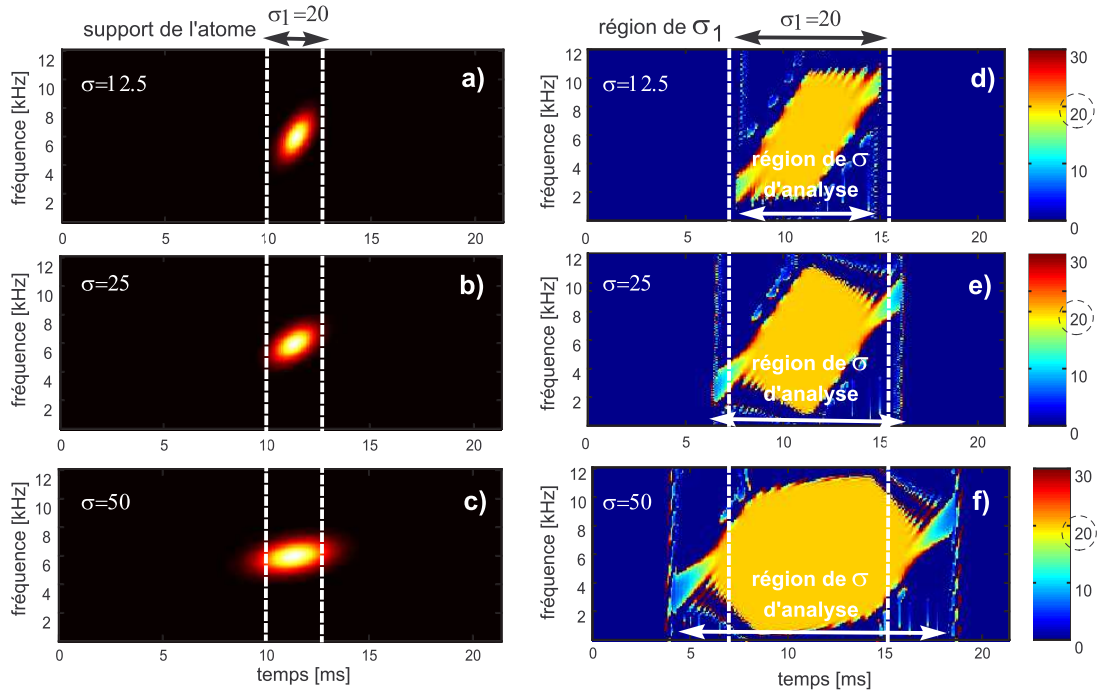


FIG. 10.8 – Un atome chirp : spectrogramme de Gabor (à gauche) et cartographie du support σ_1 (à droite) avec une fenêtre d'analyse $\sigma = 12,5$ (en haut), $\sigma = 25$ (au milieu) et $\sigma = 50$ (en bas)

En revanche, dans le cas d'un signal constitué de deux et plusieurs atomes, ces perturbations aux bords peuvent influencer la lisibilité de la représentation. D'une manière analogue, on peut définir la cartographie du paramètre β_1 qu'on va appeler cartographie de l'orientation de l'atome. Elle ressemble à la cartographie du support et permet aussi de détecter les limites entre les composantes du signal.

Exemple 3 : Cartographies du support et de l'orientation de deux atomes chirp Les cartographies du support et de l'orientation de deux atomes chirp fournissent les valeurs estimées de σ_k et β_k pour chaque point temps-fréquence. Ces cartographies indiquent la séparation des composantes du signal dans le plan temps-fréquence et donnent, entre autres, un indice permettant d'expliquer pourquoi parfois le signal reconstruit contient des atomes qui ont été positionnés différemment des atomes du signal initial. En effet, dans certains cas (notamment quand les atomes sont très proches dans le plan temps-fréquence), nous verrons que l'estimation peut être influencée par l'interaction des supports des atomes.

Pour illustrer cette situation, on prend deux signaux, chacun constitué de deux atomes chirp. Le premier signal analysé (Fig. 10.9 a, c, e, g) est constitué de deux atomes proches avec le même paramètre $\sigma_1 = \sigma_2 = 20$ et des paramètres de chirp différents $\beta_1 = 0.005$ et $\beta_2 = -0.005$. Le spectrogramme de Gabor avec une fenêtre d'analyse $\sigma = 25$ (Fig. 10.9 a) montre que les supports des deux atomes interagissent. Cette interaction provoque l'apparition d'un troisième point, différent des "centres" des deux atomes. Finalement, on verra plus loin que le signal reconstruit comporte deux atomes, dont le deuxième est positionné justement au point d'intersection des atomes. Par ailleurs, le point d'intersection possède une grande énergie, due à la superposition des deux atomes chirp (Fig. 10.9 e). La cartographie des orientations montre également le point d'intersection des atomes qui est situé à mi-distance entre les centres des deux atomes (Fig. 10.9 g).

Le deuxième signal analysé (Fig. 10.9 b, d, f, h) comprend deux atomes possédant le même paramètre de chirp, mais dont les supports sont suffisamment éloignés dans le plan temps-fréquence (Fig. 10.9 b). Dans ce cas, les atomes sont bien séparés et leurs supports n'interagissent pas (Fig. 10.9 d). Cependant, on peut apercevoir sur la cartographie du support une variation des valeurs des points situés autour de la ligne de séparation des supports, en formant une zone, qu'on peut désigner comme instable (Fig. 10.9 f). On peut supposer qu'un rapprochement des supports des atomes modifiera davantage les valeurs du support σ_k dans la zone instable et influencera l'estimation des paramètres. D'une manière générale, les cartographies du support et de l'orientation (paramètres σ_k et β_k) donnent une indication du comportement numérique de la sélection des atomes, décrite en détail dans la section suivante. Elles

peuvent indiquer pourquoi lors de la sélection un point temps-fréquence est privilégié devant un autre.

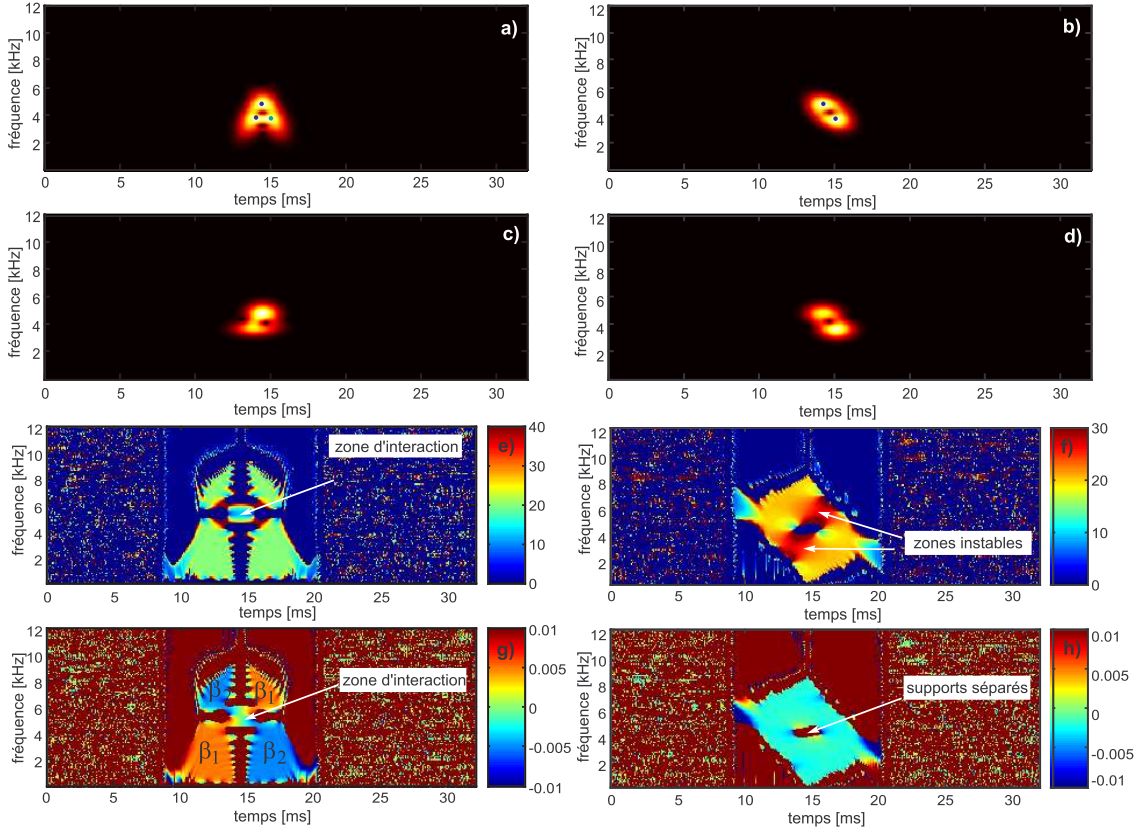


FIG. 10.9 – Le spectrogramme du signal original (a,b) et le signal reconstruit de deux atomes chirp (c,d). Cartographie des supports (e,f) et des orientations (g,h) des atomes.

Une situation plus "extrême" comprend deux atomes chirp localisés au même point (t_1, ω_1) du plan temps-fréquence. Les deux atomes ont des supports et des paramètres de chirp différents. Les cartographies du support et de l'orientation témoignent que les paramètres des deux atomes sont déterminés d'une manière unique seulement hors de la zone d'intersection des deux supports. Les valeurs dans la zone d'intersection ne sont pas vraiment prévisibles, car les deux atomes interagissent trop fortement.

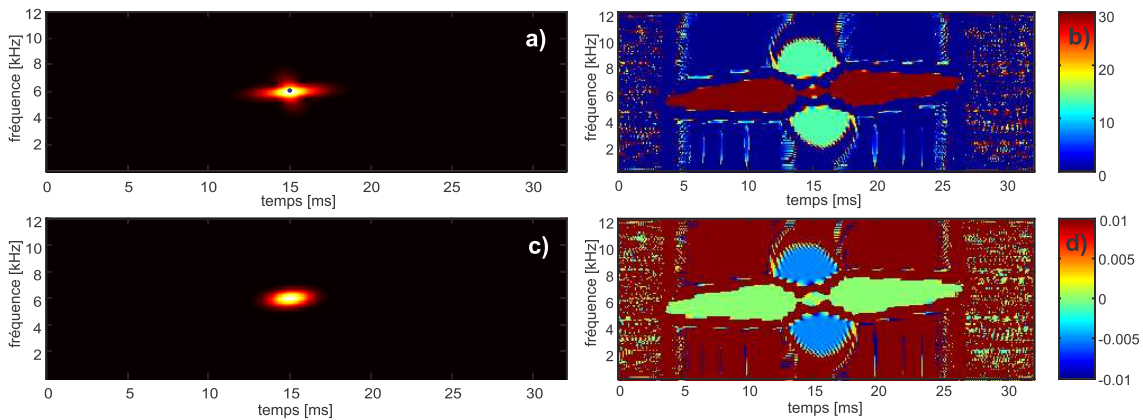


FIG. 10.10 – Spectrogrammes de Gabor de deux atomes positionnés au même point, le signal original (a) et le résultat de notre algorithme (c) Cartographie de σ_1 et σ_2 (b), β_1 et β_2 (d)

Le point choisi est également le point le plus énergétique, situé justement au centre d'intersection des deux atomes. Pour la reconstruction du signal sont prises les valeurs de σ_1 et β_1 situées dans le point d'intersection. Sur la **Fig. 10.10**, c'est l'atome "horizontal" qui est finalement retenu. Des situations

analogues laissent supposer qu'il serait judicieux d'estimer les paramètres hors de la zone d'intersection des atomes, en tenant compte du support sur lequel σ_1 où β_1 sont définis et restent constants.

10.3 Algorithme d'estimation des paramètres

La grande redondance de la famille d'atomes de Gabor nécessite une sélection des atomes. Notre première sélection consiste à extraire les maxima locaux du module de la transformée de Gabor donnant un premier ensemble de points candidats. Un deuxième critère mesure l'indépendance entre les atomes dans le plan temps-fréquence afin d'éviter leur recouvrement. Un troisième critère assure que le support estimé des atomes σ_k n'est pas trop grand ou trop petit par rapport au support σ de la fenêtre d'analyse.

10.3.1 Maxima locaux

Une façon de réduire la redondance de la famille d'atomes de Gabor consiste à prendre seulement les points où est concentrée la plus grande partie de l'énergie du signal. Un ensemble de points importants est constitué par les maxima locaux du module de la transformée de Gabor.

Motivation de l'approche Le choix des maxima locaux est une approche qui est au cœur des méthodes de type "arête et squelette" où les lignes d'"arête" dans le plan temps-fréquence sont constituées des maxima locaux de la transformée de Gabor [71]. Par ailleurs, il est montré pour l'algorithme de "ridge pursuit" que le meilleur atome chirp peut être estimé à partir de l'ensemble des maxima locaux de la projection du signal sur un sous-ensemble du dictionnaire de Gabor. En effet, les maxima locaux contiennent l'information relative à la fréquence instantanée et à ses variations [127]. Jusqu'à présent, la fréquence instantanée a presque toujours été reliée à la phase et non au module de la transformée de Gabor. Cependant, il est déduit que la phase peut être entièrement déterminée par le module et vice-versa, en démontrant que le spectrogramme porte autant d'informations que la transformée à valeur complexe [114]. Cet argument accompagne le fait que tous les vecteurs de la réallocation du spectrogramme dans le cas d'une fenêtre gaussienne, pointent dans la direction des maxima locaux du module. Le traitement du spectrogramme se rapproche du traitement des images [175].

Calcul direct à partir du module Le maximum local est un point dont la valeur est maximale, comparée aux valeurs des points de son voisinage. Le moyen le plus courant de calculer les maxima locaux est issu directement de la définition. Il s'agit de vérifier, pour chaque point (t_i, ω_j) du plan temps-fréquence, que la valeur du module de la transformée de Gabor en ce point est supérieure ou égale à la valeur dans tous les points du voisinage (**Fig. 10.11**). Ceci implique 8 conditions vérifiées à la fois :

$$\begin{aligned} M(t_i, \omega_j) &\geq M(t_{i+1}, \omega_j), & M(t_i, \omega_j) &\geq M(t_{i+1}, \omega_{j+1}), \\ M(t_i, \omega_j) &\geq M(t_{i-1}, \omega_j), & M(t_i, \omega_j) &\geq M(t_{i-1}, \omega_{j+1}), \\ M(t_i, \omega_j) &\geq M(t_i, \omega_{j+1}), & M(t_i, \omega_j) &\geq M(t_{i+1}, \omega_{j-1}), \\ M(t_i, \omega_j) &\geq M(t_i, \omega_{j-1}), & M(t_i, \omega_j) &\geq M(t_{i-1}, \omega_{j-1}), \end{aligned}$$

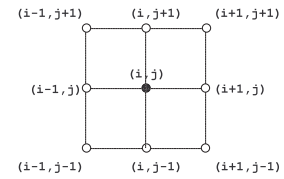


Fig. 10.11 – Le voisinage du point (i, j)

Cette approche utilise le voisinage le plus proche du point (t_i, ω_j) qui, en pratique, est discrétisé. On pourrait éventuellement élargir le voisinage au niveau supérieur, mais ceci impliquerait encore 16 conditions supplémentaires. Cette méthode directe extrait très bien les maxima locaux, cependant elle a parfois tendance de choisir deux maxima très proches (voir les exemples plus loin). Ceci n'engendre pas une grande différence dans la sélection finale des atomes, car une sélection par la suite, appliquée sur le critère de distance entre les atomes, élimine les atomes considérés trop proches. Il est nécessaire d'appliquer le critère de proximité des atomes pour éliminer les maxima trop proches, car l'algorithme de réallocation risque de réallouer les deux maxima au même point, en superposant ainsi deux atomes à la place d'un seul.

Points fixes de la phase Un autre moyen pour estimer les maxima locaux repose sur la recherche des maxima locaux des fonctions de la fréquence locale et du retard local. Ceci est exprimé à partir des dérivées de la phase de la transformée de Gabor. Les points fixes sur la ligne de la fréquence locale sont détectés dans un point (i, j) si les deux points adjacents $(i, j + 1)$ et $(i, j - 1)$ sont de signe opposé. L'approche utilisée dans le contexte des méthodes d'"arête et squelette" pose deux conditions pour la

détection des maxima locaux sur les lignes spectrales [71] :

$$\frac{\partial \Phi}{\partial \tau}(i, j) - \omega(i) \geq 0 \quad \text{et} \quad \frac{\partial \Phi}{\partial \tau}(i+1, j) - \omega(i+1) \leq 0. \quad (10.35)$$

D'une manière similaire, à partir des lignes du retard du groupe, les conditions sont

$$\frac{\partial \Phi}{\partial \alpha}(i, j) \geq 0 \quad \text{et} \quad \frac{\partial \Phi}{\partial \alpha}(i, j+1) \leq 0. \quad (10.36)$$

Dans le cas d'un atome, les lignes spectrales et du retard de groupe se croisent en un point commun, le centre de l'atome. Ainsi, pour la détection des maxima locaux d'un signal constitué d'atomes temps-fréquence, les quatre conditions (Eq. 10.35) et (Eq. 10.36) doivent être satisfaites.

Dans tous les cas, quelque soit la méthode choisie pour calculer les maxima locaux, il est pratique de faire un tri rapide des points en ordre décroissant. En fixant une limite supérieure l_{max} , et la plus grande valeur du module M_{max} , on sélectionne seulement les points dont le module est suffisamment grand, i.e

$$(i, j) : \quad M_{(i, \omega_j)} > \frac{M_{max}}{l_{max}}. \quad (10.37)$$

Ainsi, on élimine des points qui représentent des maxima locaux, mais possèdent une énergie très faible.

10.3.2 Critère de distance

Les atomes sélectionnés par le critère des maxima locaux peuvent se retrouver très "proches" entre eux. Dans cette situation, il est judicieux de garder un seul à la place de deux ou plusieurs atomes.

Recouvrement temps-fréquence La distance entre les atomes dans le plan temps-fréquence est un problème très proche de la réallocation dans le sens où il s'agit d'attribuer correctement l'énergie aux différentes composantes constituant le signal. Ce problème est soulevé par Chassande-Mottin pour la méthode de la réallocation différentielle, et comparé à d'autres méthodes comme les partitions de Voronoï et la mesure d'entropie de Renyi [114]. Pour un signal constitué d'atomes gaussiens à variance unité, trois paramètres des atomes influencent la partition temps-fréquence : la distance entre les atomes, leurs amplitudes relatives et leurs phases. A une amplitude et une phase identiques pour deux atomes, il est observé que leur spectrogramme de Gabor avec une fenêtre à variance unité (voir Eq. 7.25) les affiche fusionnés en un seul atome si la distance entre les deux atomes devient plus petite que la valeur critique de $\sqrt{2}$. Au-dessus de la distance critique, le spectrogramme admet trois extrema, dont deux maxima situés aux "centres" des deux atomes, et sépare bien les supports des deux atomes. A une distance très proche, pour deux atomes à amplitude différente, l'atome possédant l'amplitude plus importante "empiète" sur la partie temps-fréquence de l'atome plus faible. Ceci est d'autant plus valable que la différence entre les amplitudes est grande. A une distance suffisamment grande, l'amplitude n'influence pas la partition et la ligne de coupure se situe à mi-distance entre les atomes. Il est donc important de s'assurer que les atomes sélectionnés par notre algorithme ne se recouvrent pas dans le plan temps-fréquence.

Produit scalaire entre deux atomes Une mesure du degré de recouvrement (ou d'indépendance) entre les atomes dans le plan temps-fréquence est donnée par le produit scalaire entre deux atomes. La formule analytique pour le calcul du produit entre deux atomes gaussiens chirpés $g_1(t)$ et $g_2(t)$ est [117]

$$\langle g_1(t), g_2(t) \rangle = \frac{1}{\sqrt{\pi\sigma_1\sigma_2}} \sqrt{\frac{\pi\bar{\sigma}^2}{\sqrt{\mu}}} e^{i\frac{\theta}{2}} e^{R+iI} \quad (10.38)$$

où plusieurs notations sont utilisées pour simplifier l'écriture

$$\begin{aligned} R &= -\frac{\Delta\tau^2}{2(\sigma_1^2 + \sigma_2^2)} - \frac{\bar{\sigma}^2}{4\mu}(\Delta\omega - \bar{\xi}_R)^2, \\ I &= \frac{(\omega_1 + \omega_2)\Delta\tau}{2} + \frac{\Delta\beta\Delta\omega^2}{8} + \frac{\Delta\beta\Delta\omega^2}{8\mu} \left(\frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2 + \sigma_1^2} \right)^2 + \frac{\Delta\tau(\Delta\omega - \bar{\omega}_I)}{2\mu} \frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2 + \sigma_1^2} - \frac{\bar{\sigma}^4\Delta\beta}{8\mu}(\Delta\omega - \bar{\xi}_I)^2, \\ \Delta\tau &= \tau_1 - \tau_2, \quad \Delta\beta = \beta_1 - \beta_2, \quad \Delta\omega = \omega_1 - \omega_2, \quad \theta = \arctan\left(\frac{\bar{\sigma}^2\Delta\beta}{2}\right), \\ \bar{\sigma}^2 &= \frac{2\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}, \quad \bar{\xi}_I = \frac{\beta_1 + \beta_2}{2}\tau, \quad \bar{\xi}_R = \frac{\beta_1\sigma_1^2 + \beta_2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\Delta\tau, \quad \mu = 1 + \frac{(\bar{\sigma}^2\Delta\beta)^2}{4}. \end{aligned}$$

Plus le produit scalaire est proche de zéro, plus les atomes sont indépendants entre eux dans le plan temps-fréquence. Le seuil, déterminé par le produit scalaire, est situé dans l'intervalle $[0,1]$ et doit être choisi expérimentalement en fonction du signal analysé. Dans les exemples numériques sur le bruit moteur, on examinera les résultats de l'analyse-synthèse avec de différentes valeurs du seuil déterminé par le produit scalaire.

10.3.3 Rapport entre σ et σ_k

Le troisième critère de sélection est lié au rapport entre σ de la fenêtre d'analyse et le paramètre estimé σ_k de l'atome. L'utilité de ce critère est d'assurer numériquement, que la valeur estimée du support σ_k est en cohérence avec la valeur σ d'analyse. Cette cohérence peut être exprimée comme une limite supérieure et inférieure pour les valeurs acceptables de σ_k . Plus particulièrement, la limite L est définie d'une manière arbitraire, pour fixer le rapport entre σ et σ_k . Elle exige que σ_k estimé soit dans les limites

$$L < \frac{\sigma}{\sigma_k} < \frac{1}{L}. \quad (10.39)$$

Ainsi, les supports σ_k trop grands ou trop petits par rapport à la fenêtre d'analyse ne vont pas être retenus. Ce critère de sélection est incorporé à deux fois dans l'algorithme : il fait partie du critère de sélection des maxima locaux et du critère du produit scalaire. Toutefois, le critère du rapport entre σ et σ_k est le moins restrictif et il n'élimine pas beaucoup d'atomes par rapport aux autres critères. Ce critère a le sens d'un contrôle du bon déroulement numérique de l'algorithme.

10.3.4 Algorithme itératif

L'estimation des paramètres a permis de formuler les nouveaux opérateurs de réallocation (Eq. 10.30 et Eq. 10.31). Initialement, les positions des atomes sont estimées sur une grille, déterminée par la discrétisation en temps et en fréquence. Or, il est possible de s'affranchir des limites de l'échantillonnage temps-fréquence et d'affiner davantage ces positions. L'affinement des positions des atomes finalement retenus est effectué à l'aide d'un algorithme itératif de point fixe, emprunté aux méthodes d'"arête et squelette". Ce processus comporte plusieurs itérations, où les valeurs de τ et α dans les formules (10.30) et (10.31) sont remplacées à chaque fois avec les nouvelles coordonnées de l'atome i , trouvées à l'itération précédente. La convergence se fera simultanément pour les coordonnées de tous les atomes (t_i, ω_i) pour $i = 1..N$. Le processus itératif pour l'itération $n + 1$ est résumé avec les formules

$$t_i^{n+1} = t_i^n - \frac{\partial \Phi}{\partial \alpha} \left(1 + \frac{\sigma_i^2}{\sigma^2} \right) + \beta_i \sigma_i^2 \sigma^2 \left(\frac{\partial \Phi}{\partial \tau} - \omega_i^n \right) \quad (10.40)$$

$$\omega_i^{n+1} = \frac{\partial \Phi}{\partial \tau} + \left(\frac{\partial \Phi}{\partial \tau} - \omega_i^n \right) \left(\beta_i^2 \sigma_i^2 \sigma^2 + \frac{\sigma^2}{\sigma_i^2} \right) - \beta_i \frac{\sigma_i^2}{\sigma^2} \frac{\partial \Phi}{\partial \alpha} \quad (10.41)$$

A chaque itération doivent être également estimées à nouveau les valeurs de σ_i et β_i de chaque atome g_i , pour $i = 1..N$. La convergence vers la "vraie" position temps-fréquence de l'atome dépend de la position initiale. Comme on le verra plus loin, si la position initiale se trouve dans un point dit "instable", il est possible que la convergence ne s'effectue pas correctement. Cependant, très souvent les points instables possèdent peu d'énergie et sont donc préalablement éliminés selon le premier critère énergétique (Eq. 10.37). Pour un atome chirp, cet ajustement se fait très rapidement et ne nécessite qu'une ou deux itérations pour réallouer les valeurs exactes de l'atome. Cet exemple confirme l'exactitude des formules analytiques dans le cas d'un seul atome chirp.

10.3.5 Estimation de l'amplitude A_k

Le paramètre qui est estimé en dernier lieu est l'amplitude de l'atome A_k . En effet, il est précédé non seulement de l'estimation de tous les autres paramètres de l'atome, mais également des procédures de sélection et d'itération, qui viennent d'être présentées.

Dans le cas d'un seul atome, l'amplitude A_1 peut être retrouvée à l'aide du produit scalaire entre l'atome g_1 dont tous les paramètres sont connus (sauf l'amplitude) avec la fonction de la famille de Gabor $h_{\tau,\alpha}(t_1, \omega_1)$, localisé sur le point (t_1, ω_1) , car, en effet,

$$G(t_1, \omega_1) = A_1 < g_1, h_{\tau,\alpha} > . \quad (10.42)$$

En supposant que le signal est constitué de la somme de plusieurs atomes, cette estimation est biaisée. En effet, les atomes, sont choisis de façon à ce qu'ils soient indépendants dans le plan temps-fréquence, en utilisant le critère du produit scalaire. Cependant, les atomes ne forment pas une base. Par conséquent, la projection sans perte d'information n'est pas possible. Ceci implique que les produits scalaires entre les atomes $\langle g_i, g_j \rangle$ pour $i \neq j$ ne sont pas rigoureusement nuls. Il faudrait éventuellement tenir compte de ces termes qui s'ajoutent, en utilisant par exemple une méthode fondée sur les moindres carrés.

10.3.6 Exemples d'estimation

Les différents aspects du tri des atomes en fonction des critères de sélection vont être illustrés à l'aide des signaux de base, constitués d'un ou de deux atomes.

Exemple 4 Le signal constitué d'un seul atome, présenté dans l'Exemple 1, est reconstruit presque parfaitement par l'algorithme d'analyse-synthèse (Fig. 10.12 b), où l'erreur numérique de l'ordre de 10^{-6} , correspond à l'erreur de troncature de la gaussienne (Fig. 10.12 c). Cette bonne estimation est due au fait que le signal utilisé reflète parfaitement le modèle atomique du signal (Eq. 10.1). Les coordonnées en temps et en fréquence de l'atome sont données d'une manière exacte par les formules analytiques et par conséquent les critères de sélection et l'algorithme itératif n'ont pas lieu d'être.

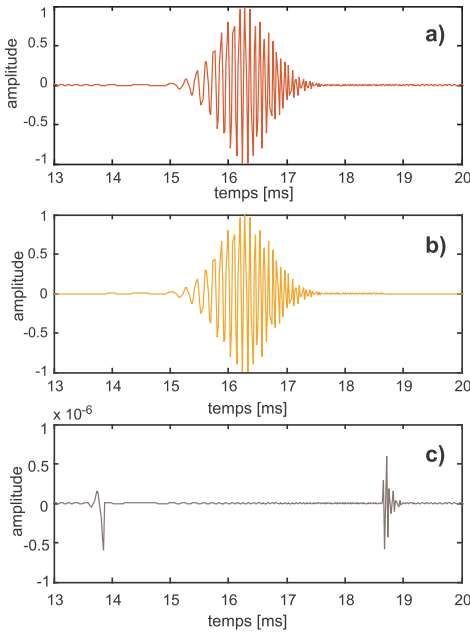


Fig. 10.12 – L'enveloppe temporelle de l'atome chirp (a), de l'atome chirp reconstruit (b) et l'erreur d'estimation (c)

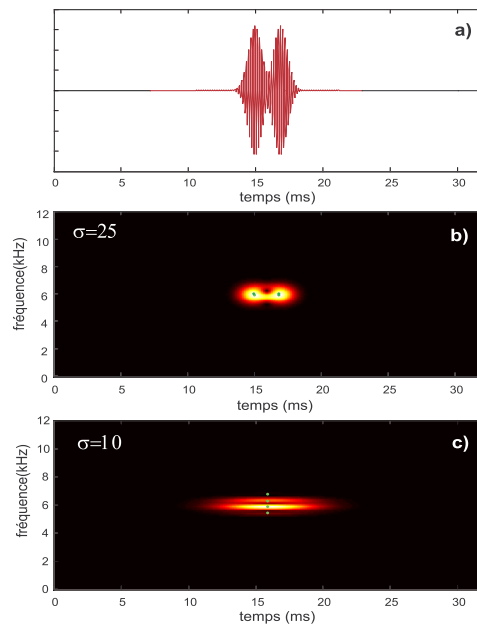


Fig. 10.13 – Deux atomes proches en temps (a). Spectrogramme de Gabor et maxima locaux détectés en fonction de la fenêtre d'analyse (b,c)

Exemple 5 : Influence de la fenêtre d'analyse L'influence de la fenêtre d'analyse sur l'extraction des maxima locaux est illustrée à l'aide d'un signal constitué de deux atomes non chirpés, proches en temps et de paramètre de support $\sigma_1 = 20$. La fenêtre d'analyse donne une visualisation différente des composantes du signal sur le spectrogramme de Gabor et influence l'extraction des maxima locaux (Fig. 10.13). Pour une fenêtre, plutôt proche de l'étendue de l'atome ($\sigma = 25$), il y a une bonne localisation et séparation des supports des deux atomes. Deux maxima locaux sont estimés. En revanche, pour une fenêtre localisée en fréquence ($\sigma = 10$), quatre maxima locaux, disposés "verticalement", sont détectés. La détection de plusieurs maxima laisse supposer que l'énergie de la fenêtre d'analyse contribue grandement à la répartition énergétique. La position en temps des maxima est située à mi-distance des instants d'apparition en temps des atomes. Par ailleurs, la valeur $\sigma = 25$ de la fenêtre d'analyse semble la plus adaptée pour la détection des atomes de courte durée (de quelques millisecondes). Le choix de cette valeur sera confirmé dans l'exemple de la synthèse d'un cycle moteur au ralenti.

Exemple 6 : Algorithme itératif à partir de maximum local Cet exemple utilise deux signaux constitués chacun d'un atome avec coordonnées en temps et en fréquence $t_1 = 16,281\text{ms}$ et $f_1 = 5512,5\text{Hz}$ et dont les étendues respectives sont $\sigma_1 = 25$ et $\sigma_2 = 10$. Le paramètre de support de la fenêtre d'analyse est $\sigma_1 = 25$. L'algorithme d'analyse sélectionne d'abord les points candidats comme les maxima locaux

du module de la transformée de Gabor. En général, pour un seul atome, un seul maximum local est extrait (**Fig. 10.14 a,b**). Dans ce cas, la coordonnée en temps de la valeur itérée et du maximum local correspond à la vraie coordonnée en temps de l'atome. L'affinement de la position de l'atome se fait ensuite seulement en fréquence. La valeur en fréquence du maximum local trouvé est $f = 5555,9\text{Hz}$, tandis que la valeur réallouée ramène l'énergie exactement sur la fréquence de l'atome, en $f = 5512,5\text{Hz}$.

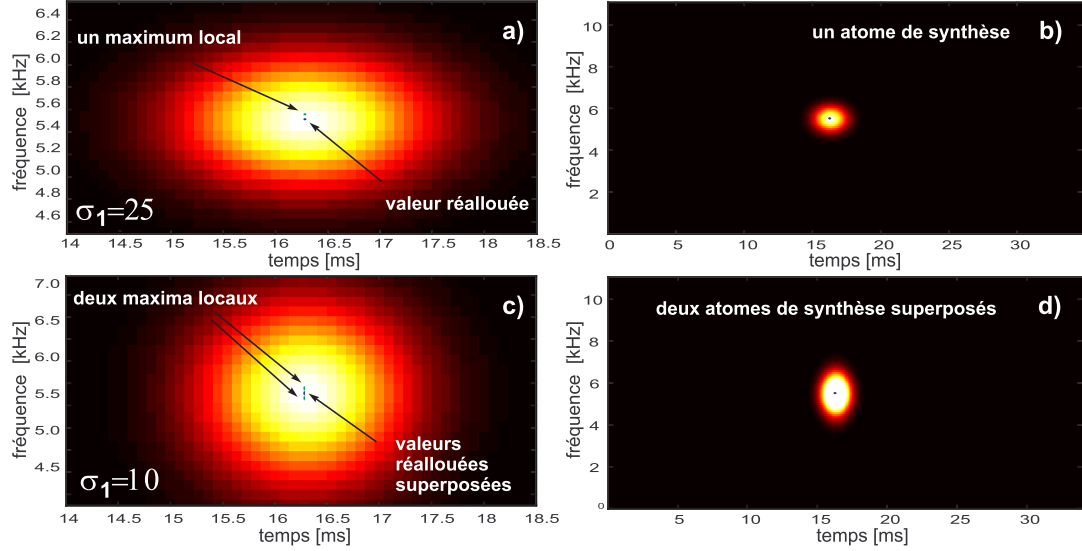


FIG. 10.14 – Itération du maximum local détecté d'un atome non chirpé en fonction de son support $\sigma_1 = 25$ (en haut) et $\sigma_1 = 10$ (en bas). En haut, un seul maximum est détecté et une seule valeur est réallouée. En bas, 2 maxima locaux sont détectés et 2 valeurs sont réallouées en un seul point

Toutefois, comme on a vu dans l'Exemple 5, la méthode d'extraction des maxima locaux parcourt un voisinage très proche et les maxima locaux trouvés pour un atome peuvent être plusieurs. Dans le cas de l'atome plus large en fréquence (moins large en temps, $\sigma_1 = 10$), deux maxima sont extraits (**Fig. 10.14 c**). Ces deux maxima sont très proches en fréquence et ont la même coordonnée en temps. Cette situation met en évidence l'utilité du critère de sélection qui retient seulement les atomes suffisamment éloignés. L'élimination est importante, car dans le cas contraire, les deux maxima locaux vont se réallouer au même point, en superposant leurs énergies (**Fig. 10.14 d**). A la place, le critère de distance retient une seule valeur qui va ensuite se réallouer sur la position temps-fréquence de l'atome.

Exemple 7 : Cartographie des supports et réallocation Le problème de la séparation des composantes du signal trop proches peut affecter également le processus de la réallocation. Dans cet exemple, on va prendre le cas d'un signal constitué de deux atomes chirp possédant la même amplitude et on va varier la distance en temps entre leurs supports. Les paramètres de chirp et de support des deux atomes sont respectivement $\beta_1 = 0.001$, $\beta_2 = 0.007$, $\sigma_1 = 20$ et $\sigma_2 = 10$.

Atomes éloignés Quand les deux atomes sont suffisamment éloignés en temps (**Fig. 10.15 a**) l'estimation des paramètres s'effectue d'une manière exacte (**Fig. 10.15 c**). Dans ce cas, les temps d'apparition des atomes, respectivement $t_1 = 14,96\text{ms}$ et $t_2 = 19,16\text{ms}$ sont suffisamment éloignés et l'estimation des paramètres $\sigma_{1,2}$ et $\beta_{1,2}$ se fait d'une manière exacte à partir des formules analytiques. De plus, les positions en temps et en fréquence n'ont pas besoin d'être affinées par l'algorithme itératif. Ceci peut s'expliquer par le fait que les expressions analytiques ont été formulées dans le cas d'un atome chirp et leur généralisation dans le cas d'un signal constitué de plusieurs atomes reste correcte seulement si les atomes sont suffisamment indépendants (i.e séparés) dans le plan temps-fréquence.

Atomes proches En revanche, dans le cas du signal constitué des mêmes atomes, mais dont les supports se touchent (**Fig. 10.15 e**), l'estimation du paramètre β_2 (et par conséquent σ_2 , puisqu'ils sont liés) est perturbée. Ici, les atomes possèdent des coordonnées en temps $t_1 = 14,96\text{ms}$ et $t_2 = 17,02\text{ms}$. Les valeurs estimées après la 5-ème itération de réallocation, sont $\sigma_1 = 19.9258$, $\sigma_2 = 16.3658$ et $\beta_1 = 0.0011$ et $\beta_2 = 0.0064$. On observe que la pente (le paramètre d'orientation) de l'atome de droite, ainsi que son support diffèrent de l'atome original. De plus, la fréquence estimée du deuxième atome est légèrement supérieure à la fréquence originale. Ceci laisse supposer que dans le cas d'un signal dont les composantes

se recouvrent dans le plan temps-fréquence, l'estimation et la réallocation se font plus difficilement. Cet exemple montre, entre autres, l'intérêt de bien sélectionner les atomes afin que leurs supports soient assez éloignés. Par ailleurs, les cartographies de σ et β des atomes (Fig. 10.15 b, d, f, h) reflètent le rapprochement ou la séparation atomes.

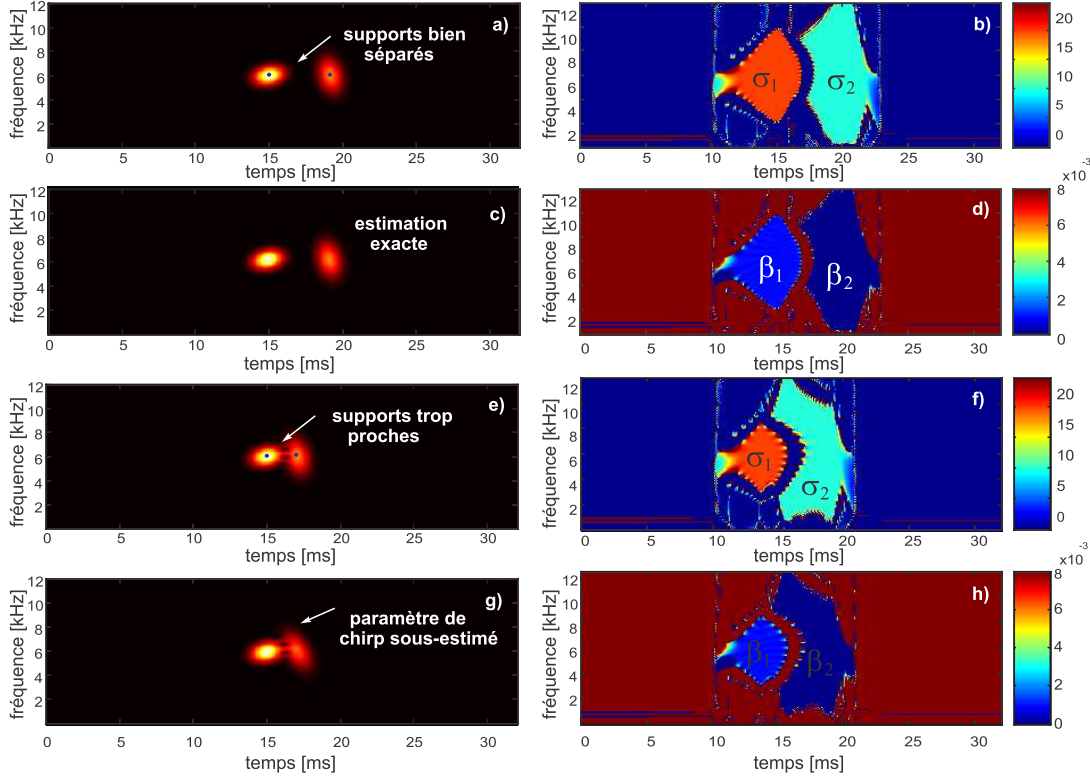


FIG. 10.15 – Deux atomes avec des supports bien séparés (a) et avec des supports proches (e). Spectrogramme des signaux originaux (a et c) et des signaux reconstruits (e et g). Cartographies des supports et des orientations : atomes éloignés (b,d) et atomes proches (f,h).

Exemple 8 : Séparation de deux atomes chirp L'estimation des paramètres des atomes peut dépendre non seulement de la distance entre les atomes, mais aussi du paramètre de chirp. Cet exemple présente l'analyse-synthèse de huit signaux : les quatre premiers signaux sont constitués de deux atomes avec la même pente ; les quatre autres signaux comportent deux atomes avec des pentes de signe opposé. La distance en temps et en fréquence entre les atomes varie également.

Atomes éloignés Dans le cas où les deux atomes chirp sont suffisamment indépendants, l'estimation se fait d'une manière exacte (Fig. 10.16 a,a'). Avec la diminution de la distance entre les atomes, que ce soit en temps ou en fréquence, il est possible jusqu'à une certaine limite que les supports des atomes soient estimés séparément (Fig. 10.16 b,b',c,c'). Dans le cas où le paramètre de chirp est modifié, l'estimation des paramètres peut être altérée, comme le montrent les signaux sur (Fig. 10.16 c,c') et (Fig. 10.16 d,d'). Dans le premier cas, les atomes ont le même paramètre de chirp et la situation est favorable pour la reconstruction correcte du signal. Dans le deuxième cas, un des atomes possède une pente différente, ce qui rapproche les supports des deux atomes, perturbe l'estimation et un seul atome est estimé. Cet atome est positionné quelque part "au milieu" de l'aire temps-fréquence couverte par les deux atomes. Les spectrogrammes des signaux reconstruits après l'algorithme d'analyse-synthèse sont marqués avec "''".

Pentes différentes Les exemples suivants concernent des signaux de deux atomes dont les pentes sont différentes. Dans le cas où les atomes sont relativement éloignés, trois maxima locaux sont initialement détectés, dont un dans le point d'intersection des supports des atomes (Fig. 10.16 e). Finalement, le maximum "parasite" n'est pas retenu, car il ne possède pas suffisamment d'énergie (Fig. 10.16 e'). Dans le cas où les atomes sont encore plus proches, le maximum au point d'intersection des supports attire une énergie plus importante (Fig. 10.16 f, g). Le résultat est que l'algorithme d'analyse-synthèse sélectionne deux maxima locaux et positionne un atome au point d'intersection (Fig. 10.16 f' et g').

Amplitude différente Quand l'amplitude d'un des atomes devient plus importante que l'autre (**Fig. 10.16 h**), la détection du deuxième atome n'a pas lieu (**Fig. 10.16 h'**). Même si dans certains cas l'estimation et la séparation des composantes ne peut pas se faire rigoureusement, le signal reconstruit présente souvent une signature temps-fréquence similaire à celle du signal original (**Fig. 10.16 f, f'**).

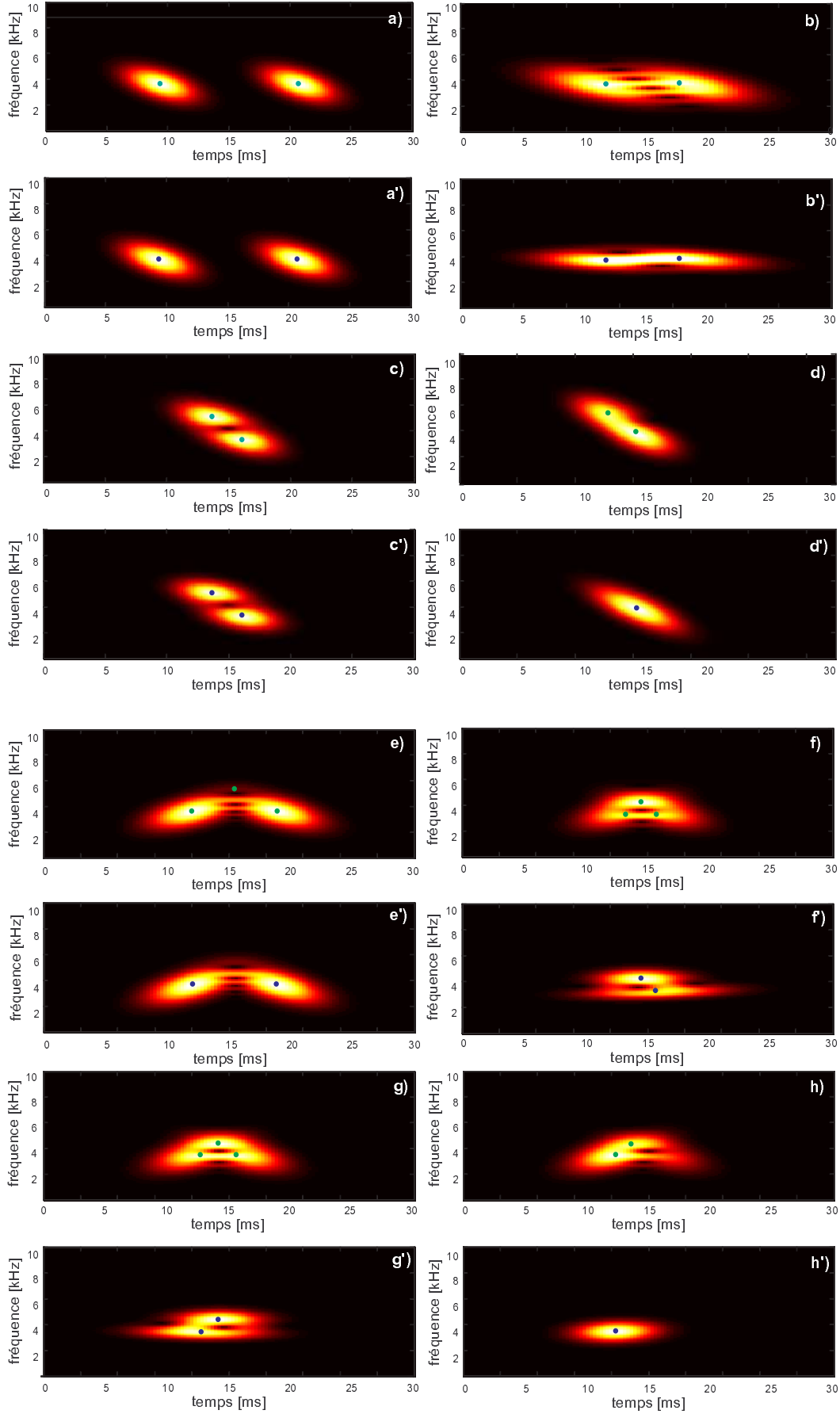


FIG. 10.16 – Estimation des paramètres de deux atomes chirp en fonction de la distance entre eux et le paramètre de chirp. Pour les détails : voir le texte ci-contre

10.4 Synthèse granulaire

La synthèse du signal s'effectue en sommant les parties réelles des atomes temps-fréquence :

$$s_{synth} = \sum_{k=1}^N Re \left\{ A_k e^{-\frac{(t-t_k)^2}{2\sigma_k^2}} e^{i[\omega_k(t-t_k) + \frac{\beta_k}{2}(t-t_k)^2]} \right\}. \quad (10.43)$$

Synthèse du bruit moteur L'application de la méthode pour les bruits moteur considère l'analyse-synthèse d'un seul cycle moteur. Le cycle moteur représente une unité qui porte, en principe, l'information complète du fonctionnement moteur. Cette restriction est adoptée aussi pour limiter le temps de calcul. A partir de l'estimation des paramètres d'un cycle moteur, on peut obtenir par périodisation de chaque atome, le bruit moteur final. Pour ce faire, il est nécessaire de fabriquer un signal suffisamment long, en enchaînant la répétition du cycle de synthèse. Ensuite, on positionne chaque atome d'une manière continue à une distance en temps égale à un cycle moteur (**Fig. 10.17**). En général, il faut périodiser les atomes sur au moins 10-20 cycles moteur, pour que l'oreille puisse assimiler le bruit de synthèse.

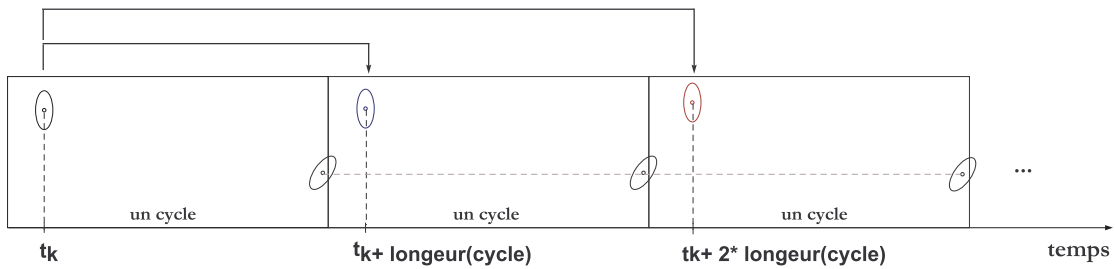


FIG. 10.17 – Principe de synthèse granulaire du bruit moteur sans discontinuités au bord du cycle

Régularité de la répétition L'inconvénient de l'analyse-synthèse sur seulement un cycle réside dans la répétition trop "régulière" du même cycle. En réalité, cette régularité est caractéristique pour les bruits impulsifs répétitifs comme le marteau-piqueur ou la mitraillette. Par rapport à ces types de sons, le bruit moteur a un caractère beaucoup plus "irrégulier", ce qui traduit les différences cycle-à-cycle. D'une manière générale, il faut analyser environ 20 cycles moteur afin de pouvoir traduire correctement cette irrégularité. Or, cette condition est trop exigeante d'un point de vue de temps de calcul.

D'un point de vue perceptif, le bruit plus irrégulier est perçu moins impulsif par l'oreille. Dans ce contexte, un bruit moteur trop régulier peut être considéré comme moins convaincant par rapport à l'original. Afin de remédier à la périodicité trop régulière, on peut analyser le bruit moteur sur plusieurs cycles (3-4 par exemple) et ensuite effectuer la synthèse avec eux. Cet exemple est présenté plus loin. Une autre idée consiste à introduire une légère modification aléatoire des atomes qui se trouvent au début et à la fin du cycle, en empêchant ainsi la répétition régulière (**Fig. 10.18**). Ces approches ne peuvent pas réduire totalement la régularité.

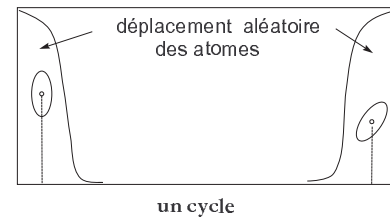


Fig. 10.18 – Modification des atomes aux bords du cycle

Toutefois, après plusieurs expériences numériques, il m'a semblé plus approprié de rester sur la synthèse d'un seul cycle moteur, tout en adaptant la fenêtre d'analyse en fonction du signal afin d'obtenir la meilleure configuration d'atomes et par conséquent, un son de synthèse plus réaliste.

Pertinence de la synthèse La méthode d'analyse-synthèse peut être appliquée sur n'importe quel signal sonore. Comme toute méthode d'analyse-synthèse, on peut se poser des questions sur la ressemblance entre le son original et le son de synthèse. Il est évident que, comme pour quasiment tous les sons de synthèse, l'oreille humaine arrive à faire la différence avec le son original. Toutefois, les sons de synthèse obtenus avec notre méthode sont suffisamment réalistes, dans la mesure où les bruits au ralenti ne sont pas remis en cause par les auditeurs. De ce fait, les sons de synthèse peuvent être utilisés pour des tests perceptifs.

10.5 Analyse-synthèse des bruits moteur

Après la description de la méthode d'analyse-synthèse et l'illustration de son application sur des signaux de base (un ou deux atomes chirps), cette section se propose de montrer l'application de la méthode pour les bruits moteur au ralenti. Les exemples numériques sont consacrés à l'extraction de la période du cycle moteur, l'estimation des paramètres, l'influence des critères de sélection, le comportement du processus itératif et enfin, la synthèse finale du bruit moteur.

10.5.1 Extraction d'un cycle

Le cycle moteur peut être facilement extrait à l'aide la fonction d'autocorrélation. Il est important de mentionner que ceci donnera une valeur moyenne de la longueur du cycle, car les bruits moteur ne sont pas, proprement dit, des signaux stationnaires. Comme on l'a déjà mentionné, il existe des variations cycle-à-cycle permettant de définir plusieurs types de cyclostationnarité [7]. Cette méthode ne sera pas utilisée ici pour des raisons d'une grande complexité d'application.

Autocorrélation et périodicité La corrélation entre deux signaux mesure leur dépendance et ressemblance réciproques. De même, l'autocorrélation, exprimant la corrélation d'un signal avec lui même, témoigne des ressemblances internes dans le signal. L'autocorrélation mesure la similitude du signal et sa version décalée d'un intervalle τ et permet d'estimer les différentes périodicités contenues dans le signal. La fonction d'autocorrélation d'un signal $s(t)$ est une fonction symétrique par rapport à l'origine et s'écrit dans le domaine discret

$$R[k] = \sum_n (s_n - m)(s_{n-k} - m), \quad (10.44)$$

où la valeur moyenne m est connue comme l'espérance mathématique.

Autocorrélation et spectre Le théorème de Wiener-Khinchin donne, via la transformée de Fourier, une relation entre la fonction d'autocorrélation et la densité spectrale. Plus précisément, elle dit que la densité spectrale d'un signal est la transformée de Fourier de la fonction d'autocorrélation qui lui correspond. La densité spectrale S_{xx}

$$S_{xx}(f) = \sum_{k=-\infty}^{\infty} R[k] e^{-i2\pi k f} \quad (10.45)$$

est une fonction symétrique et son premier maximum à partir de son centre de symétrie donne la période fondamentale du signal. Cette approche est souvent utilisée pour l'estimation de la fréquence fondamentale d'un son, puisque ceci revient à chercher une périodicité fondamentale dans le signal. La fonction d'autocorrélation est une manière simple d'avoir une estimation de la longueur du cycle moteur. Cependant, comme le bruit moteur contient les contributions de plusieurs sources, il se peut que la fonction d'autocorrélation soit influencée par les périodicités internes, en provenance d'une source dont la périodicité est double (voir les exemples numériques).

La recherche du maximum de la fonction d'autocorrélation dans un intervalle proche de la valeur moyenne du régime moteur peut faciliter l'extraction de la période fondamentale du bruit moteur.

Exemple 9 : Extraction d'un cycle La longueur du cycle moteur est égale à la distance entre le centre de symétrie et le premier maximum de la fonction d'autocorrélation. Toutefois, la fonction d'autocorrélation n'extrait pas les mêmes périodes pour des sons enregistrés à 6 endroits différents du compartiment moteur (**Fig. 10.20**). Il s'agit d'un moteur à essence, à un régime 3000 tr/min (voir l'Annexe B). La période globale de $2N$ du cycle moteur est altérée par la période $0.5 N$ de la soupape d'échappement et par la période N pour le vilebrequin. La période du son enregistré près de l'alternateur n'est pas bien déterminée. De ce fait, il est nécessaire de "guider" l'extraction du cycle, en imposant un intervalle de recherche du maximum de la fonction d'autocorrélation, en se basant sur l'information approximative du régime moteur. La longueur du cycle estimée est d'environ 35 ms.

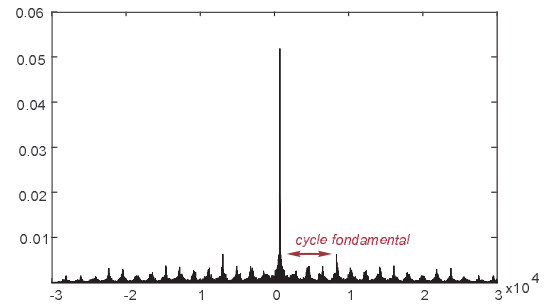


Fig. 10.19 – Autocorrélation d'un cycle moteur au ralenti

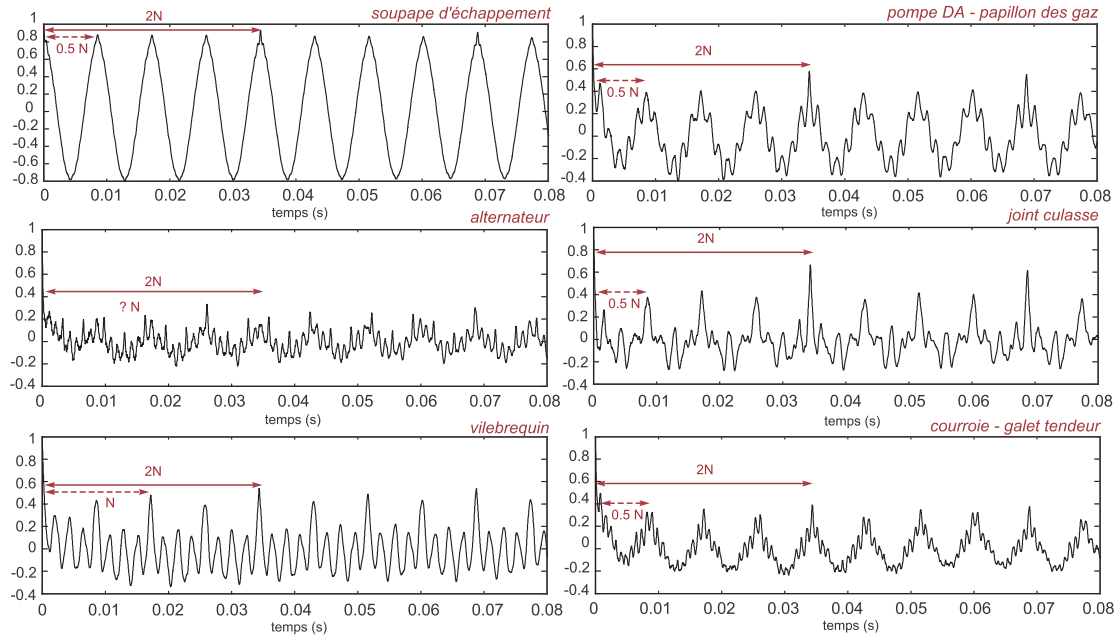


FIG. 10.20 – Fonction d'autocorrélation de sons enregistrés aux 6 endroits du compartiment moteur

La diversité des bruits dans le compartiment moteur nécessite de choisir (souvent à l'oreille) un seul bruit représentant au mieux l'ensemble du moteur. Or, cette simplification ne tient pas compte du mélange de sources différentes.

Exemple 9 : Extraction d'un cycle au ralenti

L'extraction d'un cycle moteur au ralenti comporte quelques particularités, car l'acyclisme moteur y est beaucoup plus exprimé qu'en haut régime. Au ralenti, les variations cycle-à-cycle sont beaucoup plus importantes et l'extraction de la période peut être influencée par le nombre de cycles dans le signal analysé. Parfois, il est difficile d'extraire la période fondamentale directement à partir de la fonction d'autocorrélation, car elle comporte beaucoup de variations. Un moyen de remédier à ce problème est d'utiliser la dérivée de la fonction de l'autocorrélation, qui semble donner des résultats plus nets sur la période fondamentale du cycle (Fig. 10.21).

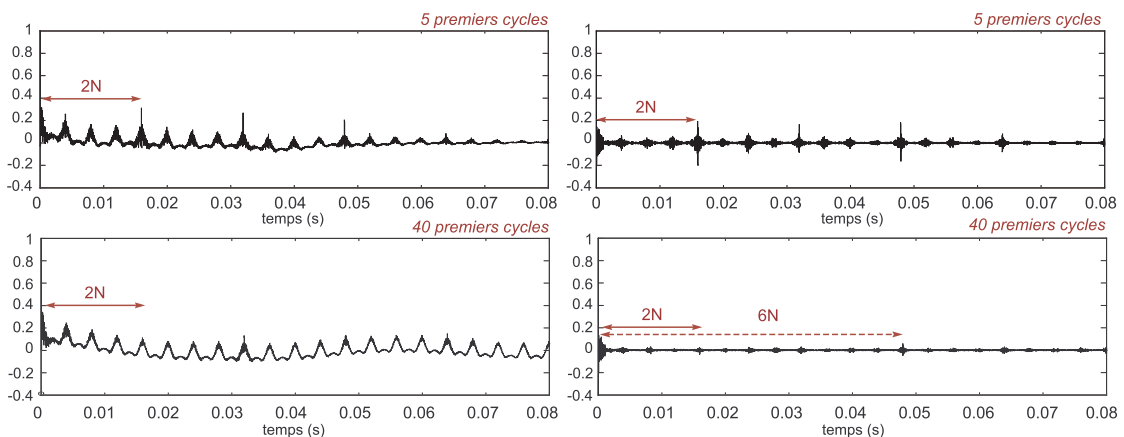


FIG. 10.21 – Bruit diesel au ralenti : fonction d'autocorrélation (à gauche) et sa dérivée (à droite)

La figure ci-dessus représente la fonction d'autocorrélation et sa dérivée, analysés sur un signal contenant respectivement 5 et 40 cycles du signal. La longueur réelle du cycle au ralenti, associée à $2N$, est d'environ 160 ms. Il semble être plus facile d'extraire le cycle à partir de la dérivée de la fonction d'autocorrélation. On aperçoit également que dans le cas d'analyse sur 40 cycles, la période globale est surestimée à $6N$. Probablement, ceci exprime un phénomène de cyclostationnarité où une période plus globale du bruit au ralenti se crée sur plusieurs cycles.

10.5.2 Estimation des paramètres

Les exemples suivants illustrent l'influence de la fenêtre d'analyse et des critères de sélection des atomes sur le résultat final de l'algorithme d'analyse-synthèse des bruits diesel au ralenti

Exemple 10 : Influence de la fenêtre d'analyse La fenêtre d'analyse détermine les propriétés de localisation en temps et en fréquence de la transformée de Gabor. Le paramètre de support σ de la fenêtre participe aux formules d'estimation des paramètres de support et d'orientation des atomes. La valeur du paramètre σ est choisi au début de l'algorithme. J'ai effectué plusieurs expériences afin de déterminer la valeur de σ la plus pertinente en vue de la synthèse sonore. Numériquement, le support discret en temps de la fenêtre gaussienne est calculé en fonction du paramètre σ , afin d'assurer une valeur minimale (de l'ordre de 10^{-5}) aux bords du support. Pour une rapidité de calcul, on prend le nombre d'échantillons le plus proche en puissance de 2, soit 256 échantillons pour une fenêtre avec $\sigma = 25$. L'influence de la fenêtre d'analyse sur le spectrogramme de Gabor est représentée à l'aide de trois valeurs de σ (**Fig.10.22**). La fenêtre avec $\sigma = 12,5$ privilégie la localisation temporelle, alors que la fenêtre avec $\sigma = 50$ est plus étendue en temps.

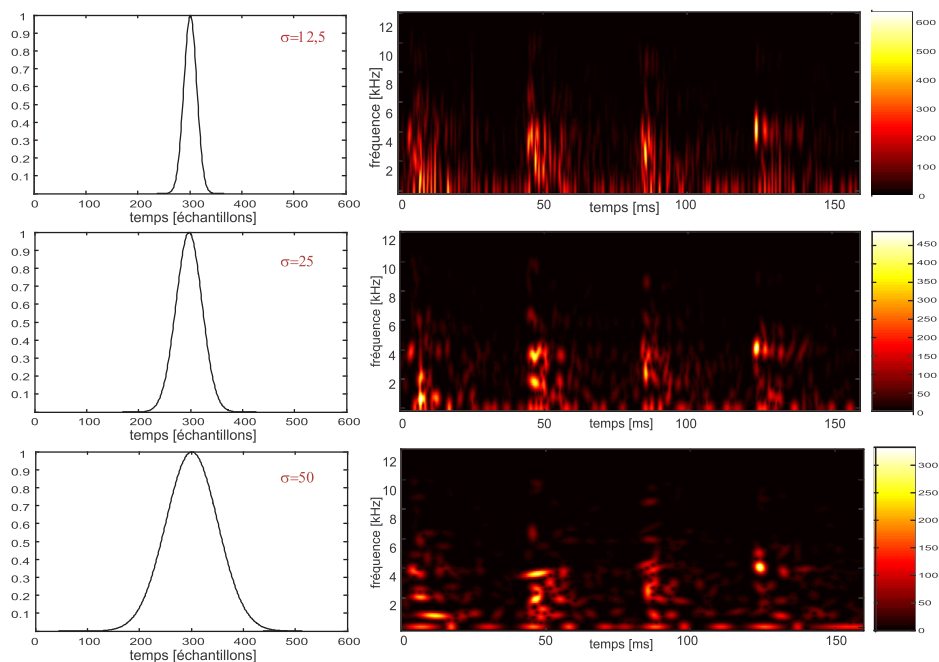


FIG. 10.22 – Un cycle au ralenti : spectrogramme de Gabor avec trois fenêtres d'analyse différentes

Des sons de synthèse sont obtenus à l'aide de l'algorithme à partir d'un bruit diesel au ralenti (**Piste 51**) et pour des valeurs respectives de la fenêtre d'analyse : $\sigma = 20$ (**Piste 52**), $\sigma = 25$ (**Piste 53**), $\sigma = 33$ (**Piste 54**), $\sigma = 35$ (**Piste 55**), $\sigma = 40$ (**Piste 56**) et $\sigma = 50$ (**Piste 57**). On peut observer qu'une fenêtre trop localisée en temps ($\sigma = 20$) rend le son de synthèse très impulsif, alors qu'une fenêtre très étendue en temps ($\sigma = 50$) le déforme spectralement. Il semble aussi que un σ "petit" privilégie les basses fréquences et un σ "grand", les hautes fréquences. Le son de synthèse le plus pertinent est obtenu avec la fenêtre d'analyse $\sigma = 25$. Après plusieurs expériences numériques et l'écoute des sons de synthèse, la valeur $\sigma = 25$ semble donner de meilleurs résultats perceptifs et est adoptée comme valeur par défaut.

Exemple 11 : Influence sur les maxima locaux L'exemple précédent met en évidence l'influence de la fenêtre d'analyse sur le résultat final de synthèse. En effet, cette influence est visible encore à l'étape de la première sélection, des maxima locaux. Contribuant dans les formules analytiques d'estimation, σ va déterminer différemment les atomes de la sélection finale. Pour trois valeurs, $\sigma = 12,5$, $\sigma = 25$ et $\sigma = 50$, il semble que la fenêtre la plus large en temps ($\sigma = 50$) extrait le plus de maxima locaux (**Fig. 10.23**). Les points sélectionnés sont indiqués par des points blancs. Il reste pour la sélection finale entre 2 et 3 fois moins d'atomes que de maxima locaux détectés. Trois exemples sonores : le bruit au ralenti original (**Piste 58**), analysé avec une fenêtre $\sigma = 25$; le bruit de synthèse avec des atomes positionnés sur 585 maxima locaux initialement détectés (**Piste 59**); et le bruit de synthèse avec la sélection finale de 220 atomes (**Piste 60**), illustrent l'influence du nombre d'atomes sur le résultat perceptif.

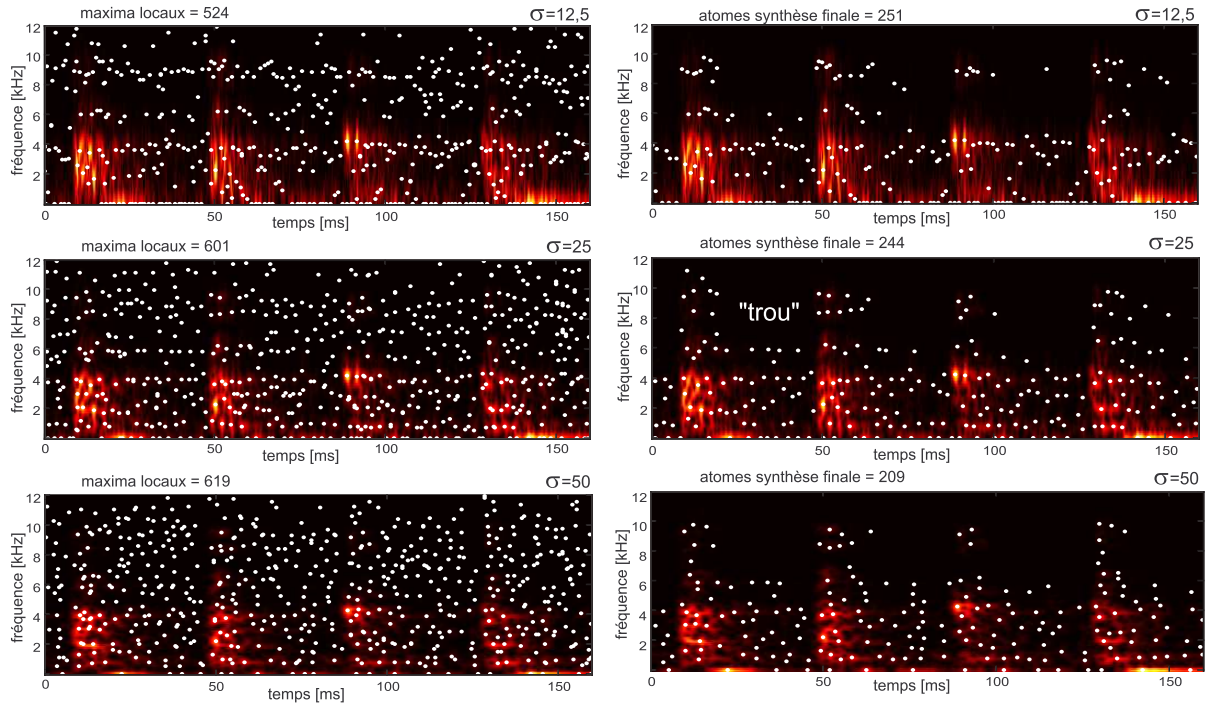


FIG. 10.23 – Spectrogramme de Gabor d'un cycle diesel au ralenti avec trois fenêtres d'analyse différentes. Les points blancs marquent les maxima locaux (à gauche) et les points finalement retenus (à droite).

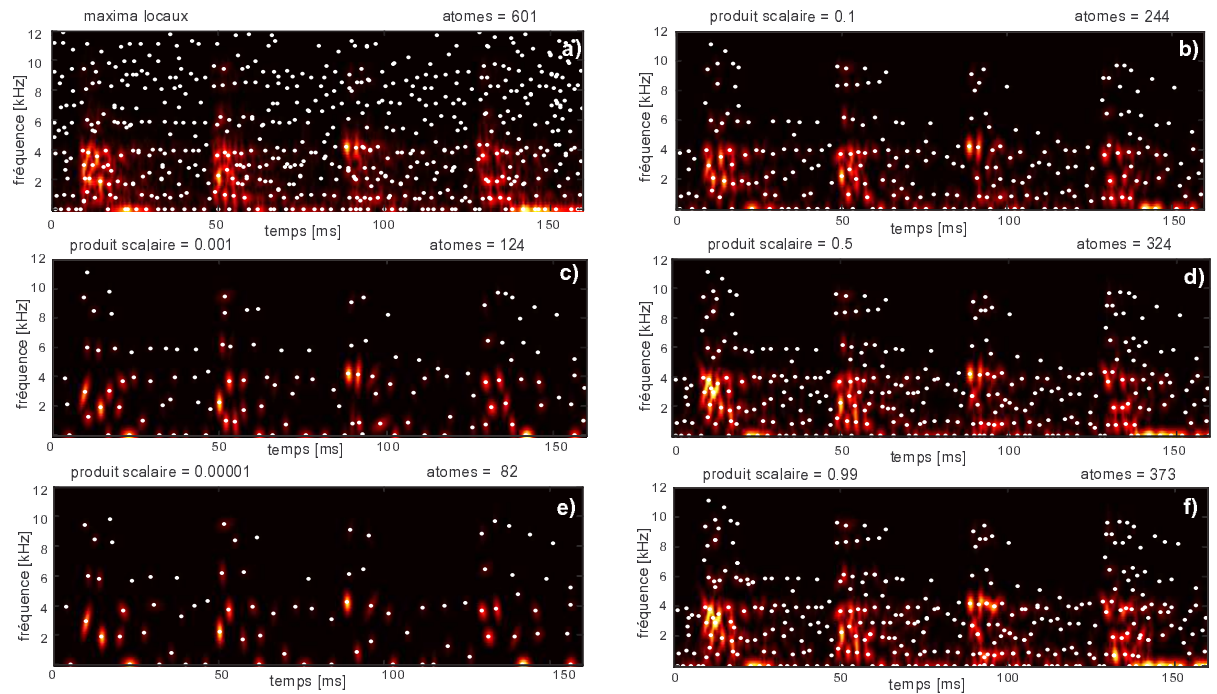


FIG. 10.24 – Indépendance entre les atomes : influence du seuil déterminé par le produit scalaire sur la sélection des atomes

Le petit nombre d'atomes a tendance à rendre le bruit de synthèse plus impulsif. En effet, en réduisant le nombre d'atomes, le flux sonore formé par les atomes possède des discontinuités (marquées par "trou" sur la Fig. 10.23), ce qui privilégie l'impulsivité du bruit. Cette observation représente un point important de notre expérience perceptive, présentée dans la troisième partie.

Exemple 12 : Critère de distance L'exemple 11 décrit l'influence de la fenêtre d'analyse sur la sélection des maxima locaux et sur la sélection finale. Cet exemple mettra en évidence l'influence de l'autre critère de sélection, la distance entre les atomes, mesurant leur indépendance. L'analyse avec une fenêtre de paramètre $\sigma = 25$ entraîne une sélection de 601 points de maxima locaux (**Fig. 10.23**). Ce nombre semble élevé, d'autant plus qu'un grand nombre des maxima locaux détectés dans les hautes fréquences, portent peu d'énergie. Généralement, les études de la combustion considèrent les hautes fréquences au dessus de 5-6 kHz sans importance acoustique et le spectrogramme des bruits au ralenti affiche une très faible énergie du signal dans la région 6-10 kHz. Toutefois, cette information peut avoir un impact perceptif, comme le montrent les sons de l'exemple 11. Ici, on va réduire successivement le nombre d'atomes afin d'apercevoir les effets sur les sons de synthèse. D'abord, on effectue un tri en ordre décroissant des maxima locaux et on prend en compte seulement les points dont l'énergie est au-dessus d'un seuil. Ce tri élimine environ la moitié des maxima locaux. Par ailleurs, on peut constater qu'une partie des points se retrouvent très proches dans le plan temps-fréquence. On a déjà vu (exemple 6) que si on n'élimine pas les maxima locaux trop proches, ceci risque de réallouer leurs valeurs en un même point, en superposant l'énergie de deux atomes à la place d'un seul. Pour éviter cette situation, on sélectionne les points selon un critère de distance (proximité). La mesure de distance est le produit scalaire entre les deux atomes.

Coefficient d'indépendance Afin d'étudier l'influence du critère de proximité entre les atomes, on fait varier le coefficient d'indépendance des atomes, exprimé par le produit scalaire. Ce coefficient est compris entre 0 et 1 et les deux limites correspondent respectivement au cas d'atomes indépendants (dont le produit scalaire est négligeable) et au cas d'atomes situés à la même position. Pour différentes valeurs du produit scalaire, on obtient le nombre suivant d'atomes situés au-dessous de ce seuil :

Coef. produit scalaire	0.00001	0.0001	0.001	0.05	0.1	0.25	0.5	0.75	0.8	0.99
Nombre d'atomes	82	99	124	214	244	284	324	344	347	373

FIG. 10.25 – Tableau de correspondance entre le seuil du produit scalaire fixé et le nombre d'atomes retenus par l'algorithme

La courbe qui correspond à la relation entre la valeur du produit scalaire et le nombre d'atomes qui se trouvent au-dessous de ce seuil est présentée sur la **Fig. 10.26**. La valeur maximale du produit scalaire (0.99) correspond à 373 atomes. Le nombre de maxima locaux détectés est de 601 atomes (**Fig. 10.23**). La différence de 228 atomes correspond au nombre de maxima locaux qui n'ont pas été retenus à cause de leur faible énergie. A la limite inférieure, on compte 82 atomes et ce nombre semble insuffisant pour la synthèse d'un point de vue perceptif. Le compromis entre un nombre restreint d'atomes et la pertinence perceptive des sons de synthèse semble atteint pour le seuil correspondant à une valeur du produit scalaire de 0.1. De manière générale, il semble convenable de choisir la valeur du produit scalaire entre 0.1 et 0.5, ce qui entraîne en moyenne un nombre total de 300 atomes par cycle.

Les différences dans la sélection des atomes en fonction du seuil de produit scalaire sont montrées sur la (**Fig. 10.24**). Cinq valeurs du produit scalaire : 0.00001 ; 0.001 ; 0.1 ; 0.5 et 0.99 sont comparées au nombre total des maxima locaux détectés (**Fig. 10.24 a**), avant l'élimination des maxima de faible énergie. Deux des valeurs du produit scalaire (**Fig. 10.24 c,e**) prises au-dessous du seuil "de compromis", 0.1 (**Fig. 10.24 b**), entraînent un nombre plus petit d'atomes sélectionnés. Les deux autres valeurs prises au-dessus du seuil (**Fig. 10.24 d,f**) donnent un nombre plus élevé d'atomes finalement retenus. Les exemples (**Fig. 10.24 e, f**) montrent les limites inférieure et supérieure de sélection. La différence entre (**Fig. 10.24 a**) et (**Fig. 10.24 f**) témoigne du nombre de maxima locaux enlevés sur le critère de faible énergie. Les sons de synthèse correspondant à un seuil de 0.5 (**Piste 61**, 324 atomes), de 0.1 (**Piste 62**, 244 atomes), de 0.001 (**Piste 63**, 124 atomes) et de 0.00001 (**Piste 64**, 82 atomes) montrent les différences perceptives et l'augmentation de l'impulsivité avec la réduction du nombre d'atomes.

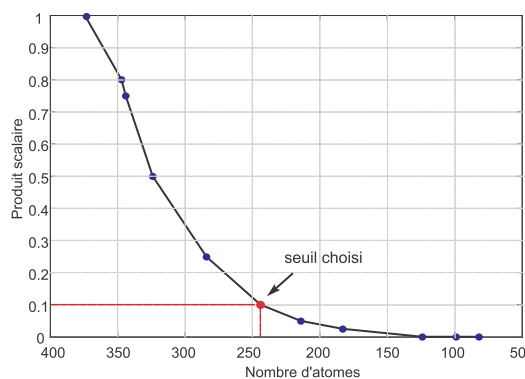


Fig. 10.26 – Critère de distance entre les atomes. Le nombre d'atomes retenus en fonction du seuil de produit scalaire entre les atomes

Exemple 13 : Cartographie de support et d'orientation de bruit moteur au ralenti La cartographie des supports et des orientations des atomes dans le cas du bruit moteur au ralenti représente une image beaucoup plus complexe que dans le cas simple d'un ou deux atomes. De ce fait, ces images sont plus difficilement interprétables. Les spectrogrammes de Gabor du signal original d'un cycle au ralenti (**Fig. 10.27 a**) du signal d'un quart de cycle (**Fig. 10.27 b**), et leur signaux de synthèse respectifs (**Fig. 10.27 c,d**) sont complétés par les cartographies du support et l'orientation des atomes. Sur les cartographies, on détecte la présence de plusieurs couleurs (valeurs) (**Fig. 10.27 f**). Ceci peut s'interpréter comme des interactions entre deux ou plusieurs atomes "proches" comme montré dans l'exemple 3.

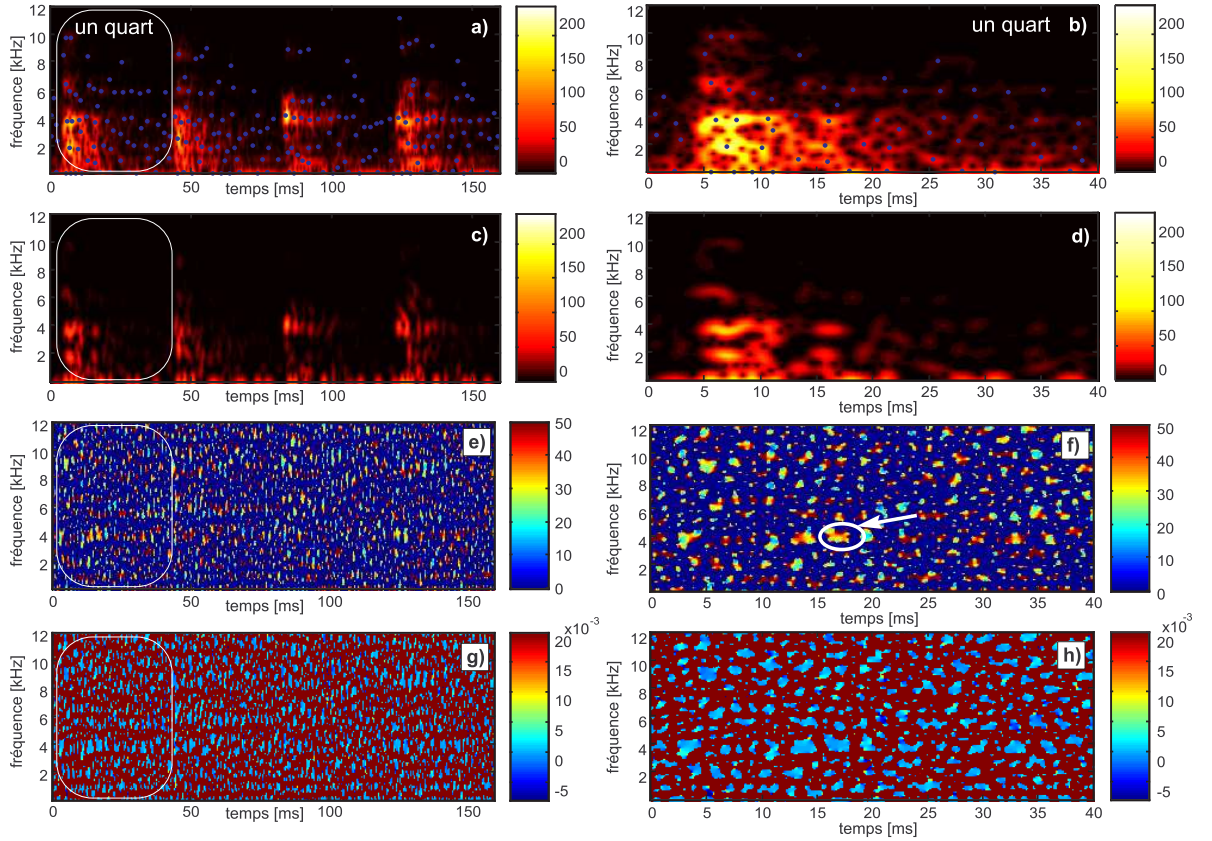


FIG. 10.27 – Bruit moteur au ralenti : un cycle (à gauche), un quart de cycle (à droite). Du haut en bas : spectrogramme de Gabor du signal original avec des points finalement retenus ; le signal reconstruit ; la cartographie de support et la cartographie de l'orientation

10.5.3 Algorithme itératif

Exemple 14 : Convergence de l'algorithme itératif L'algorithme itératif ne peut pas assurer rigoureusement la convergence vers les vraies coordonnées de l'atome à partir d'un point arbitrairement choisi. La convergence dépend de l'endroit (stable ou instable) où se trouve la position initiale du point réalloué et du nombre d'itérations effectuées. Un endroit "stable" est souvent un point qui possède beaucoup d'énergie par rapport à son voisinage. L'élimination des atomes de faible énergie est donc justifiée.

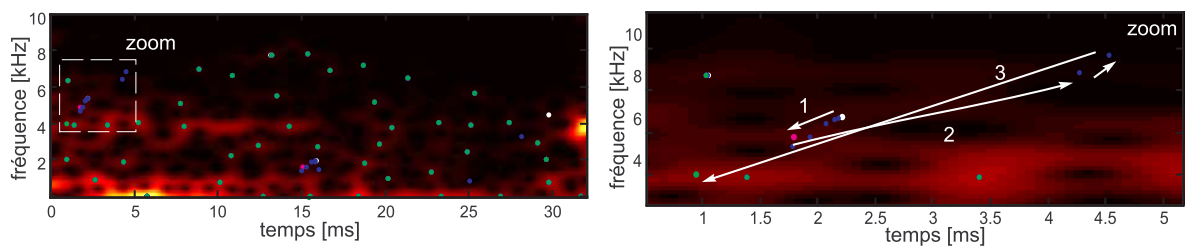


FIG. 10.28 – Un quart de cycle moteur au ralenti et la réallocation d'un point instable

Une situation où le point réalloué s'avère instable est montrée sur (**Fig. 10.28**). Le zoom sur une portion de quelques ms du cycle illustre la réallocation d'un point instable. On effectue 10 itérations au total. La position de départ est marquée par un point blanc. Les points réalloués suivants sont en bleu (trajet de réallocation marqué avec flèche "1"), alors que le point obtenu après la 5e itération est en rose. On peut apercevoir par la suite une grande divergence vers deux points situés vers 4.5 ms (trajet "2"), et le point se réalloue finalement en un point situé autour de 1 ms (trajet "3"). Ce cas montre que, à partir de la cinquième itération, l'algorithme itératif diverge. Le point tente de se réallouer sur des positions qui sont plus énergétiques et sont, en quelque sorte, des "attracteurs" d'énergie. En règle générale, les points stables se réallouent très rapidement, au plus tard à la 5e itération.

Divergence de points Pour le cycle moteur au ralenti, il existe très peu d'atomes (environ 5 dans l'exemple de **Fig 10.29**) qui ne convergent pas vers des points très proches de leur position temps-fréquence initiale. Sur la (**Fig 10.29 a**), les positions temps-fréquence après la 10e et dernière itération sont marqués en vert, alors que les positions initiales sont en blanc et les positions à la 5e itération, en rose. Toutes les autres itérations intermédiaires sont en bleu. Le fait qu'on voit essentiellement des points verts témoigne que la réallocation des positions temps-fréquence de ces atomes s'est faite dans un voisinage très proche de la position initiale. Les couleurs différentes des points permettent de suivre le déplacement de la position de ces quelques points "égares" qui divergent lors de l'algorithme de réallocation. Les déplacements sont tracés à l'aide d'un zoom sur trois zones : "zoom 1", "zoom 2" et "zoom 3".

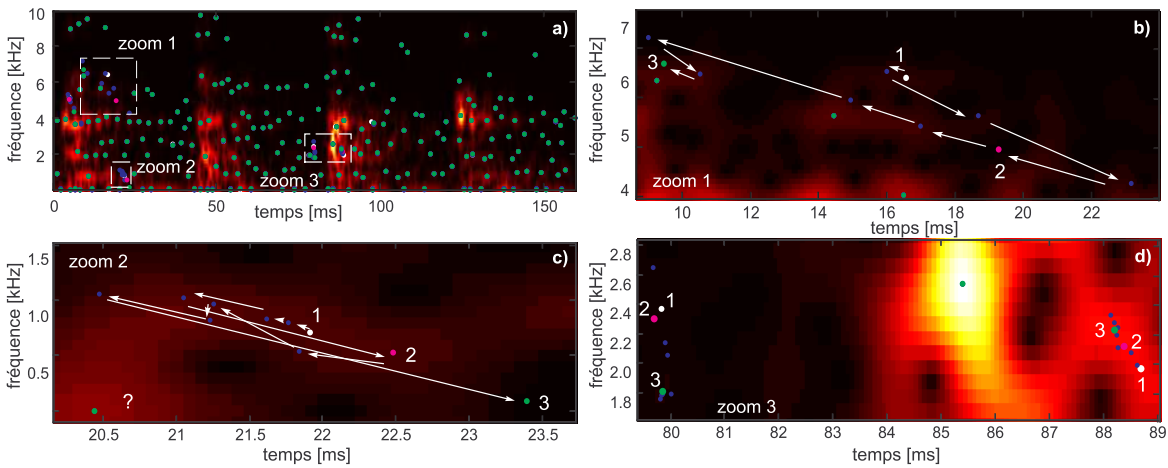


FIG. 10.29 – Divergence de certains points instables lors de l'algorithme itératif de réallocation

Les flèches indiquent un sens de déplacement un peu aléatoire pour certains de ces points. Il semble qu'ils tendent à se réallouer sur des points plus énergétiques. Toutefois, l'exemple "zoom 2" montre un point qui finit sa réallocation dans une zone de très faible énergie. En tout cas, les points qui ne convergent pas possèdent une énergie assez faible. Au contraire, les points qui ont beaucoup d'énergie se réallouent très rapidement ("zoom 3"). Ceci permet de conclure que cinq itérations sont suffisantes pour réallouer les positions des atomes et que les quelques points divergents, à cause de leur faible énergie, ne sont pas d'une grande importance pour la synthèse finale.

10.5.4 Synthèse

Trois exemples vont illustrer la synthèse des bruits moteur : la synthèse sur plusieurs cycles, le déplacement aléatoire des atomes aux bords du cycle et la synthèse du bruit moteur à haut régime.

Exemple 7 : Synthèse sur plusieurs cycles La synthèse du bruit moteur à partir d'analyse d'un cycle moteur se fait en sommant les parties réelles des atomes finalement retenus, i.e en répétant chaque atome avec une période égale à un cycle moteur. Cependant, la synthèse sur seulement un cycle engendre un effet de régularité et de répétition qui semble un peu "dur" à l'oreille. Une étude comparative des bruits de synthèse obtenus à partir de l'analyse sur un, deux et quatre cycles moteur montre que, malgré les similarités dans les signatures sonores des explosions dans chaque cylindre, les différences cycle-à-cycle sont clairement présentes (**Fig. 10.30**). Les sons de synthèse, obtenus à partir du son original (**Piste 58**), et correspondant à la synthèse d'un (**Piste 65**), deux (**Piste 66**), trois (**Piste 67**), quatre (**Piste 68**)

et dix cycles (**Piste 69**) confirment le fait que plus le nombre de cycles analysé est élevé, plus le son de synthèse semble naturel. Les particularités des cylindres sont assez bien présentées par la synthèse. Dans le cas où la synthèse est faite à partir d'un ou deux cycles, on entend un bruit répétitif assez régulier, alors que la synthèse à partir de 4 cycles semble moins régulière. La synthèse sur 3 cycles semble être la plus naturelle, probablement grâce à une répétition d'un nombre impair de cycles.

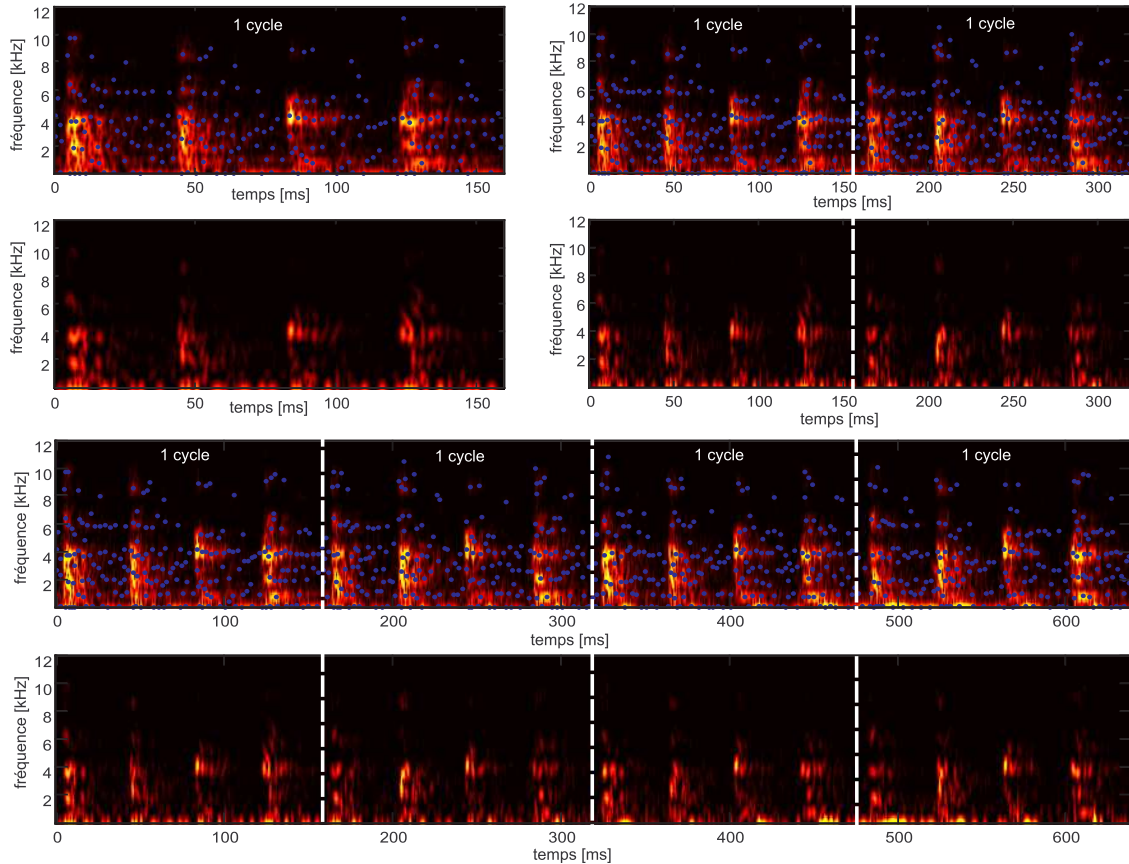


FIG. 10.30 – Le spectrogramme du signal original et reconstruit d'un, deux et quatre cycles moteur diesel au ralenti

Il semble que la régularité est liée étroitement à la notion d'impulsivité, car moins le bruit moteur semble régulier, moins il semble impulsif. Par ailleurs, la synthèse sur 10 cycles n'apporte qu'un faible changement perceptif. Le son de synthèse, obtenu de la répétition sur 3 cycles moteur analysés suggère que l'analyse-synthèse sur un nombre impair de cycles pourrait s'avérer mieux adaptée que sur un nombre pair.

Exemple 15 : Ajout de l'aléatoire La deuxième suggestion pour atténuer la répétition trop régulière consiste à modifier d'une façon aléatoire des positions des atomes au début et à la fin de la portion signal analysé. Cette portion peut contenir un, deux ou plusieurs cycles moteur. Trois exemples sonores illustrent cette modification sur un seul cycle. Le premier représente un son de synthèse sur un cycle sans modification aléatoire (**Piste 65**). Le deuxième son possède une modification aléatoire dans une région au début et à la fin du cycle qui représente 10% du cycle total (**Piste 70**). La modification s'exprime sur la position en temps des atomes. Le troisième son possède une modification du temps des atomes dans une région 30 % du cycle total (**Piste 71**) est devient trop irrégulier. En résumé, l'ajout d'une modification aléatoire peut influencer la synthèse, mais ce changement doit aussi être contrôlé.

Exemple 16 : Synthèse d'un bruit moteur à haut régime La méthode d'analyse-synthèse permet aussi la synthèse d'un bruit moteur en haut régime (**Fig.10.31 a,b**). Les sons original et de synthèse sont sur (**Piste 72**) et (**Piste 73**). La hauteur du bruit est bien restituée, mais le résultat sonore est moins convaincant que sur les bruits au ralenti. Le son de synthèse du bruit moteur dans l'habitacle (**Fig. 10.31 c,d**) montre que la méthode est moins performante, car elle garde toujours un léger aspect impulsif dans le bruit (**Piste 75**), alors qu'en réalité les bruits dans l'habitacle, surtout à un régime élevé, sont plus feutrés et ne sont pas impulsifs à l'oreille (**Piste 74**).

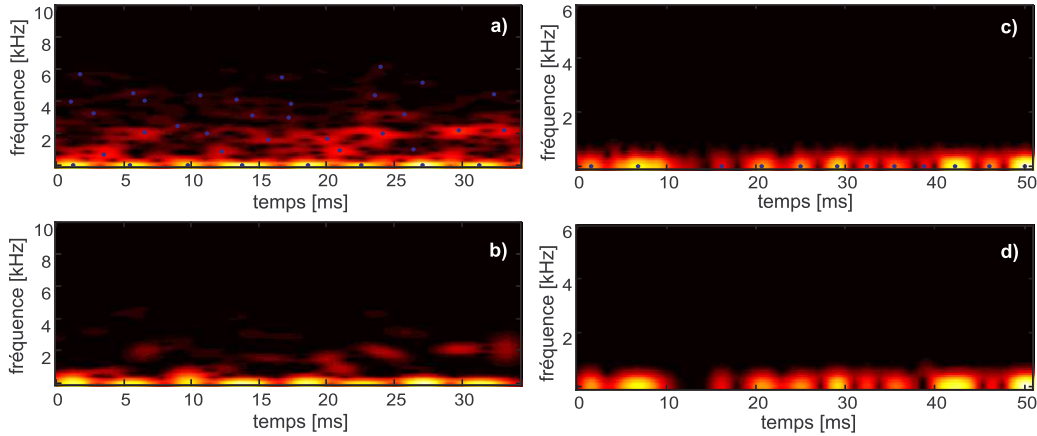


FIG. 10.31 – Spectrogramme d'un cycle moteur à 3000 tr/min dans le compartiment moteur : signal original (a) et de synthèse (b). Analyse-synthèse du bruit à haut régime dans l'habitacle : signal original (c) et de synthèse (d)

Ces exemples montrent les points faibles de la méthode. Elle ne semble pas très bien adaptée au traitement des sons non impulsifs. Probablement, la périodisation d'un cycle moteur en haut régime n'est pas une bonne solution pour la synthèse du bruit total. En haut régime on a plutôt une perception d'un bruit continu, stationnaire. Par ailleurs, le bruit dans l'habitacle possède une énergie nettement plus faible et un spectre moins étendu que le bruit du moteur dans le compartiment. Un autre point important est le nombre restreint d'atomes pour la synthèse qui ne peut pas reproduire fidèlement les nuances dans le bruit de fond qui sont présentes dans le bruit enregistré dans l'habitacle. Toutes ces observations permettent de se rendre compte des limites du modèle. Toutefois, la méthode reste pertinente pour la synthèse des bruits moteur au ralenti.

Exemple 17 : Synthèse de bruits moteur diesel au ralenti Cet exemple conclut le chapitre d'analyse-synthèse des bruits au ralenti et la deuxième partie temps-fréquence. Deux bruits au ralenti, sans injection pilote (**Fig. 10.32 a**) et avec injection pilote (**Fig. 10.32 b**) sont analysés. Ils se trouvent respectivement sur **Piste 107** et **Piste 109**. Les spectrogrammes des sons de synthèse respectifs (**Fig. 10.32 c,d**) montrent que la méthode d'analyse-synthèse reconstitue la différence entre les impulsivités des sons originaux. En l'occurrence, le son avec injection pilote est moins impulsif que celui sans injection pilote.

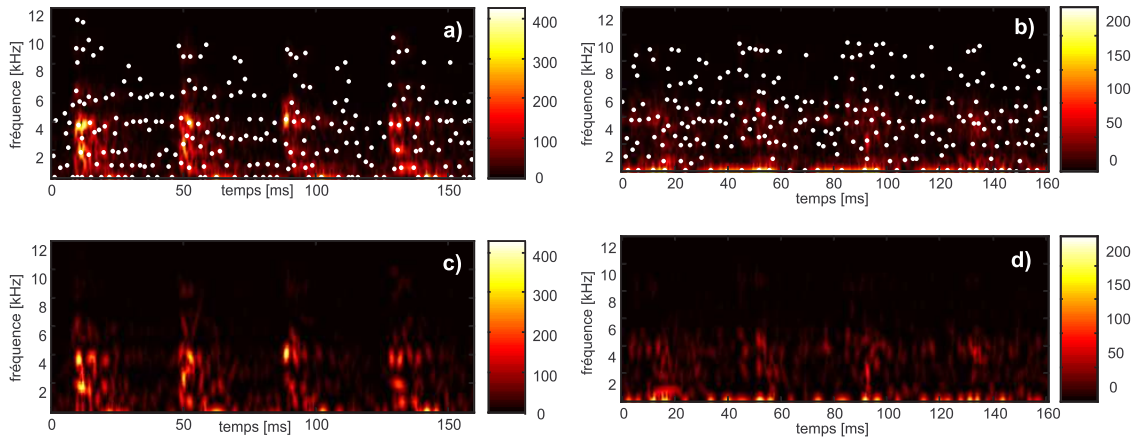


FIG. 10.32 – Spectrogrammes de deux bruits au ralenti : les signaux originaux (en haut), les signaux de synthèse (en bas). Avec des points blancs : les atomes finalement retenus pour la synthèse.

Les deux signaux de synthèse (**Piste 108** et **Piste 110**) ne sont pas complètement identiques aux sons originaux, mais reflètent d'une manière satisfaisante à l'oreille les propriétés impulsives des sons et vont être utilisés dans la troisième partie de la thèse, consacrée à l'étude perceptive des bruits moteur au ralenti.

Conclusion sur la méthode d'analyse-synthèse

La méthode d'analyse-synthèse qui vient d'être décrite a été développée dans la thèse. Elle repose sur un modèle granulaire du signal et consiste à présenter le signal comme une somme d'atomes temps-fréquence. Les points forts de la méthode sont : la possibilité d'estimation directe des paramètres des atomes à l'aide des formules analytiques ; une nouvelle définition des opérateurs de la réallocation qui permettent à concentrer l'énergie du signal sur des points dans le plan temps-fréquence plutôt que sur des lignes ; un nombre restreint d'atomes issu des critères de sélection ; la possibilité de modifier directement les paramètres des atomes temps-fréquence. Les points faibles de la méthodes sont : la difficulté d'analyser des signaux d'une durée très longue ; la régularité qui apparaît lors de la périodisation du signal de synthèse d'un cycle moteur ; la difficulté de prévoir le déroulement de tous les critères de sélection et le processus itératif.

Orientation de la thèse : La méthode d'analyse-synthèse qui est orientée vers l'analyse et la synthèse de tout signal, mais dans la thèse elle a été appliquée en particulier sur les bruits moteur au ralenti. En effet, le modèle granulaire semble plus approprié pour l'analyse-synthèse des bruits au ralenti moteur, car il permet une modélisation plus adaptée des impulsions que la synthèse additive. En revanche, même si à l'aide de la méthode, l'analyse-synthèse des bruits moteur à un haut régime reste tout à fait réalisable, les exemples sonores montrent que le modèle atomique garde une impulsivité des bruits en haut régime beaucoup trop importante que ce qu'elle est en réalité. Ce fait suggère que probablement la nature des bruits moteur et la perception auditive on peut distinguer deux mécanismes : l'un plutôt spectral, l'autre plutôt temporel. Le résultat principal de cette partie est l'algorithme d'analyse-synthèse développé. Il permet une approche originale temps-fréquence dans l'objectif d'explorer certains mécanismes de perception, propres aux bruits moteur au ralenti. L'étude psychoacoustique fait l'objet de la troisième partie de la thèse.

Troisième partie

Perception des bruits moteur au ralenti

Résumé de la troisième partie

La troisième partie de la thèse est consacrée à l'étude des bruits de moteur au ralenti dans le contexte des scènes auditives. C'est une approche innovante dans le contexte des bruits moteur. J'ai divisé la troisième partie en deux grands chapitres :

L'analyse des scènes auditives fait l'objet du premier chapitre. J'ai inclus ici des concepts sur lesquels repose l'analyse des scènes auditives comme la théorie de Gestalt, la perception des flux auditifs, les notions de cohérence et ségrégation temporelle et spectrale, etc. Une partie de ce chapitre comprend des informations complémentaires sur le processus de l'audition comme la mémoire auditive et la notion de l'attention dans la perception sonore. Enfin, je me suis proposé de montrer l'existence de structures dans le bruit moteur au ralenti, qui peuvent être interprétées dans le contexte des scènes auditives. Un point fort de l'analyse des scènes auditives consiste à considérer la cohabitation à la fois des critères temporels et fréquentiel dans l'organisation auditive. Naturellement, notre méthode d'analyse-synthèse temps-fréquence, trouve ici une réelle application.

Les tests psychoacoustiques des bruits au ralenti représentent l'aboutissement concret de l'idée de traiter les bruits moteur comme une succession de flux auditifs temps-fréquence. A partir de deux bruits de moteur, obtenus à l'aide d'un réglage différent des paramètres de la combustion, par la synthèse granulaire sont obtenus deux bruits de synthèse. Ces deux bruits de synthèse sont à la base respectivement d'une famille de bruits, tous obtenus à l'aide des modifications appliquées localement sur les atomes temps-fréquence du signal. Deux types de tests sont effectués sur les deux familles de bruits de synthèse : test de catégorisation et test de préférence. Le test de catégorisation permet de détecter les similarités entre les sons, alors que le test de préférence témoigne de l'appréciation des sons, et ce, en fonction de la modification temps-fréquence apportée.

Analyse des scènes auditives

L'information sonore qui nous entoure, comme la musique, la parole, ou les différents bruits, n'arrive pas jusqu'au nos oreilles d'une manière isolée. Notre système auditif est confronté à la tâche de traiter à la fois plusieurs événements sonores. Comment le système auditif organise cette information ? Cette question repose à la base de l'analyse des scènes auditives. La présentation des bruits moteur au ralenti dans le contexte des scènes auditives est, à ma connaissance, une idée très peu exploitée. Cette idée est toutefois justifiée, car le bruit moteur représente une superposition de multiples flux sonores, chacun en provenance de différentes sources dans le moteur. Cette nouvelle approche du bruit au ralenti mettra en rapport la théorie des flux auditifs et l'étude perceptive des bruits impulsifs répétitifs. La perception des flux auditifs est influencée par des facteurs à la fois temporels et spectraux. Cette vision s'inscrit dans la démarche d'analyse-synthèse temps-fréquence exposée dans la deuxième partie de la thèse. L'objectif de ce chapitre est de présenter les principes des scènes auditives dans la perception des flux auditifs et de relier ces expériences dans le contexte des bruits moteur au ralenti.

11.1 Principes de l'analyse des scènes auditives

Les sons dans la nature arrivent d'une façon mixte à nos oreilles. Ces événements se produisent simultanément ou séparément en temps, proviennent de différents endroits dans l'espace, et possèdent chacun un timbre spécifique caractérisant la source émettrice. L'habileté du système auditif à traiter cette information fait qu'on est capable suivre une conversation dans un milieu bruyé [176], de reconnaître le timbre de la flûte dans une pièce orchestrale, de suivre la polyrythmie des percussions africaines ou de distinguer, à partir d'un son d'impact, la forme et le matériau de l'objet frappé [96]. Cette section présente le processus de traitement de l'information sonore, la théorie de perception de la Gestalt, la perception des flux auditifs et leur interprétation pour les bruits moteur.

11.1.1 Traitement de l'information auditive

Le traitement de l'information sonore par le système auditif se fait en plusieurs étapes. Le premier processus, la transduction, assure l'acheminement des vibrations sonores vers l'oreille interne, ainsi que leur transformation en impulsions nerveuses qui progressent dans le cerveau via le nerf auditif (**Fig. 11.1**). Pour simplifier la tâche de transmission, le système auditif utilise des principes de regroupement, en associant les sons d'un groupe comme provenant de la même source. Le processus de regroupement précède la tâche d'extraction des attributs perceptifs et de présentation mentale de l'environnement sonore. Les processus de groupement auditif permettent l'organisation des événements sonores en flux auditifs, en effectuant la fusion et la ségrégation entre les sons.

Le groupement auditif se base sur des critères comme la proximité temporelle et spectrale entre les sons et précède l'étape de l'extraction des attributs perceptifs. En effet, les propriétés auditives sont déterminées en fonction du regroupement effectué. Ce procédé fait l'objet de la perception des séquences auditives. Par la suite, les attributs auditifs peuvent être interprétés à travers des structures de connaissances abstraites. Ces structures reflètent l'expérience auditive du passé, permettant à l'auditeur d'associer un son à une action. En écoutant la bande sonore d'un film, on peut s'imaginer ce qui se passe à l'écran, en associant une action avec un son particulier, par exemple l'ouverture d'une porte, le bris de verre. L'élaboration d'une représentation mentale de l'environnement sonore constitue l'étape finale du traitement des événements sonores.

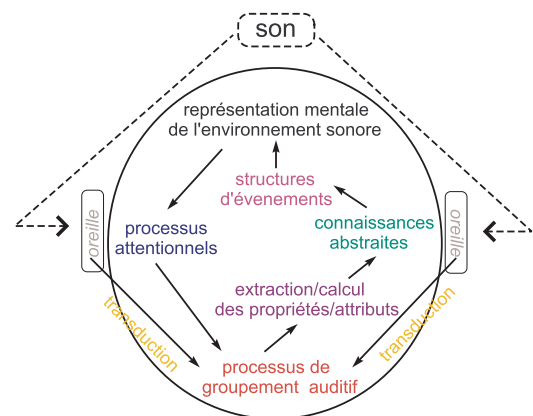


Fig. 11.1 – Principaux processus de traitement de l'information auditive (adapté de [39])

Cette étape est reliée à la mémoire auditive et permet d'anticiper la perception d'une prochaine succession de sons, ainsi que d'influencer les processus attentionnels de l'écoute [39].

11.1.2 Théorie de Gestalt

Le principe de regroupement des sons par le système auditif est inspiré par la théorie psychologique de Gestalt. Le mot allemand "Gestalt" se traduit en français un peu près comme "forme". Cette "théorie de la forme" est introduite par le psychologue allemand Max Wertheimer au début du XXe siècle [177].

Le concept de Gestalt L'idée principale du Gestalt postule que le "tout" est différent de la somme de ses parties. Notre perception prend en compte l'interaction entre les éléments constituant la forme et organise l'information selon des règles de regroupement, connues comme des lois de l'organisation perceptive. Parmi les grandes lois de perception du Gestalt on peut compter les lois de proximité, de similarité, de continuité, de clôture et de "destin commun" (**Fig. 11.2**). Ces principes sont adaptés pour l'audition dans le contexte des scènes auditives par Bregman [38].

Principe de proximité Ce principe postule que les sons dans une séquence sont regroupés selon leur proximité, soit en temps, soit en fréquence. Pour une séquence lente, les sons sont perçus l'un après l'autre en temps (**Fig. 11.2 a**). En revanche, pour une séquence rapide, les sons possédant la même fréquence vont former des flux sonores (**Fig. 11.2 b**). Les critères de proximité en temps et en fréquence entrent parfois "en compétition" et c'est l'attention auditive qui détermine quelle type d'écoute va être privilégiée. Cette expérience fait l'objet de la ségrégation et la fusion des flux auditifs, présentées plus loin.

Principe de similarité La règle de proximité est un cas particulier du principe plus commun, celui de la similarité. Selon la théorie de Gestalt, le regroupement est basé sur des critères de similarité entre les objets. Dans l'audition, la hauteur, la sonie, le timbre ou la localisation spatiale peuvent être des critères de similarité. Ces indices de ressemblance annoncent souvent au système auditif l'appartenance des sons à la même source sonore.

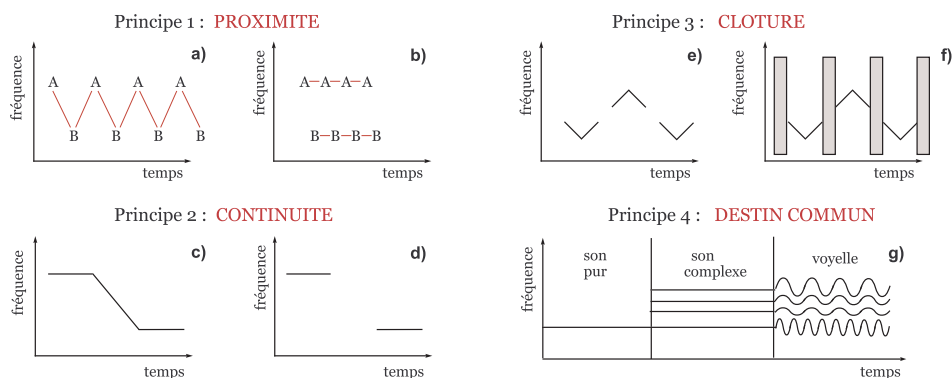


FIG. 11.2 – Les principes de Gestalt, adaptés pour l'audition par Bregman

Principe de continuité Pour que l'oreille puisse désigner les sons comme appartenant à la même source, il est nécessaire que les variations du spectre ou de la sonie soient les plus "fluides" possibles. Si des sons graves et aigus sont reliés avec des glissandi, un seul flux est alors perçu (**Fig. 11.2 c**). Dans le cas contraire, deux flux se forment, un des sons graves et un des sons aigus [178]. Le changement brusque du spectre ou de la sonie est associé au début d'un nouveau son (**Fig. 11.2 d**), en provenance d'une source différente [179]. Le principe de continuité est, en quelque sorte, le cas limite de la règle de proximité.

Principe de clôture Le principe de clôture illustre que la perception est capable de "finaliser" la forme. L'exemple donné par Bregman consiste à prendre des sons sous la forme de glissandi montants et descendants possédant de coupures (**Fig. 11.2 e**). Du bruit plus fort que les sons est introduit aux endroits des discontinuités (**Fig. 11.2 f**). Il masque les coupures et les glissandi sont perçus continus. L'exemple montre que le système auditif compense le masquage en complétant la forme auditive.

Principe du destin commun Ce principe suppose que les éléments auditifs, évoluant dans le temps de la même manière, proviennent d'une source commune. L'évolution en temps suppose un début et une fin communs des composantes sonores, ainsi que des lois similaires d'évolution en amplitude ou en fréquence. Le destin commun peut expliquer la perception des sons complexes et les voyelles. Par exemple, l'oreille associe les partiels d'un son complexe à la fréquence fondamentale du son pur lui correspondant, en considérant qu'il s'agit du même objet sonore (**Fig. 11.2 g**).

11.1.3 Perception des séquences auditives

Une grande partie des études en psychoacoustique est consacrée à la perception des sons isolés. Le premier qui met l'accent sur l'étude des séquences de sons est van Noorden [37]. Il montre que l'oreille ne traite pas les sons d'une séquence "son par son", mais utilise des mécanismes de regroupement. Van Noorden appelle cette connection entre les sons "cohérence" et constate qu'elle apparaît dans le cas d'un tempo rapide. L'intérêt de cette approche réside dans la possibilité d'expliquer plus précisément les phénomènes comme le masquage antérieur et postérieur, la modification de la sonie et de la durée perçue entre les sons dans une séquence, les processus de regroupement et des illusions auditives [180].

Ségrégation et fusion des flux auditifs La formation de flux auditifs est illustrée à l'aide de séquences construites en alternant des sons aigus de type A-A-A ... et des sons graves G-G-G ... (**Fig. 11.3**).

La perception des séquences dépend de deux paramètres : la rapidité d'exécution et l'écart fréquentiel entre les sons. Il s'agit donc d'une perception qui prend en compte à la fois les deux dimensions du son, le temps et la fréquence. A un tempo lent, et pour tout écart fréquentiel, la séquence est perçue comme une alternance de sons graves et aigus de type AGAG ..., appelée encore un trille (**Fig. 11.3 a,b**). A une vitesse rapide, la perception du trille est conservée pour un petit écart fréquentiel entre les sons (**Fig. 11.3 d**). La grande différence fréquentielle entre les sons, combinée avec un tempo rapide mène à la perception de deux flux distincts : un flux de type A-A-A .. dans les aigus, et un autre de type G-G-G .. dans les graves (**Fig. 11.3 c**).

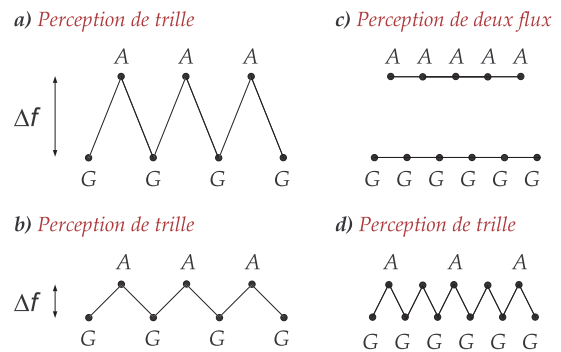


Fig. 11.3 – La formation des flux auditifs

Les flux auditifs sont le résultat de la cohérence (ou encore la fusion) entre les sons et expriment le principe de proximité des sons en temps ou en fréquence.

Etudes des flux auditifs La formation des flux auditifs est un phénomène connu et largement exploité dans la musique. Les compositeurs de l'époque du baroque du XVII-XVIII siècle comme Bach et Telemann utilisent l'alternance rapide de sons des registres aigu et grave pour créer l'illusion de plusieurs lignes mélodiques, jouées conjointement par un seul instrument (**Piste 76**). Cette technique musicale, connu comme polyphonie, fait l'objet d'une étude de Dowling [181]. Un autre exemple sur la polyrythmie dans la musique africaine est donné par Wegner [182]. La ségrégation et la fusion des séquences auditives sont explorées aussi dans le but d'estimation du seuil physiologique de ségrégation, que Miller et Heise trouvent entre 2 et 3 demitons pour la séquence AGAG ... de deux sons purs assez éloignés en fréquence et pour un tempo de 10 tons par seconde [183]. La relation temps-fréquence est confirmée par Schouten qui trouve que pour des sons présentant un grand écart fréquentiel, il est nécessaire de diminuer le tempo afin de maintenir la cohérence temporelle [184]. La sonie et le timbre influencent également la ségrégation.

Seuil de ségrégation et de fusion L'étude de la formation des flux auditifs passe par l'estimation de deux seuils : de cohérence temporelle (fusion) et de ségrégation. Van Noorden estime les deux seuils pour des séquences de types GAG ... GAG (**Fig.11.4 a,b**), où la durée des sons D est constante et égale à 40 ms et la période T est variable et permet d'accélérer ou ralentir la séquence (**Fig.11.4 c**). Les seuils de cohérence et de ségrégation sont estimés en fonction de la rapidité de la séquence, de l'ordre de 40-50 ms pour une séquence rapide et 700-800 ms pour une séquence lente, ainsi qu'en fonction de l'écart fréquentiel entre les sons, exprimé en demi-tons (**Fig.11.4 d**). Van Noorden met en évidence l'existence de deux types d'écoute dans la région située entre $T = 50ms$ et $T = 150ms$, et surtout autour de la limite critique $T = 100ms$. Physiologiquement, l'oreille possède la capacité de percevoir dans cette région à la fois la cohérence et la ségrégation. C'est grâce à l'attention que l'auditeur privilégie l'un ou l'autre type d'écoute. L'attention est donc un facteur très important dans la perception des séquences. La fusion et la ségrégation de séquences de type AGA .. AGA est illustrée par Bregman (**Piste 77**).

L'interprétation de Van Noorden de ces résultats relie le seuil de cohérence temporelle au concept temps-fréquence. Selon lui, le temps et la fréquence forment une seule relation d'un ordre supérieur déterminée par le rapport temps/fréquence. Cette approche est en accord avec l'idée de la perception auditive conjointe temps-fréquence de Gabor et les recherches des seuils de perception dans la synthèse granulaire. L'existence d'au moins deux mécanismes distincts dans la perception de la hauteur des sons complexes,

dont un spectral et l'autre temporel, a été étudiée par Grimault [185]. En revanche, ce concept temps-fréquence diffère de la théorie des canaux, où des limites des capacités du système auditif s'appliquent séparément aux dimensions du temps et de la fréquence [186].

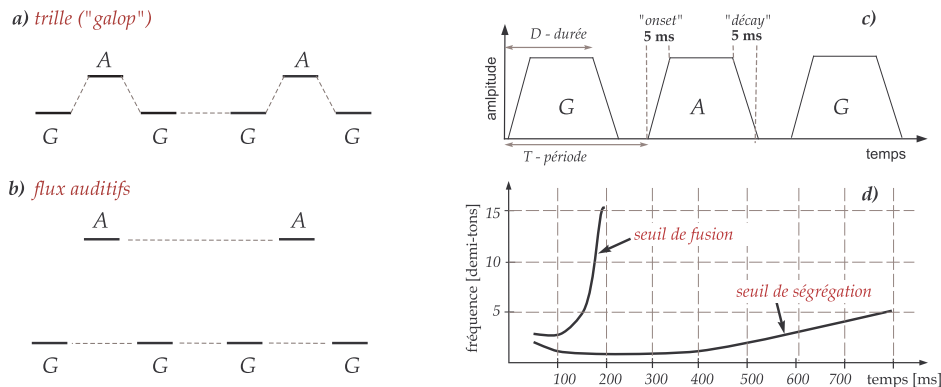


FIG. 11.4 – Ségrégation (a) et fusion (b) des séquences de sons (c). Estimation du seuil de ségrégation et de fusion par Van Noorden (d) [37].

Résolution locale et intégration globale Les séquences auditives sont étudiées aussi dans les domaines de la parole et la musique, où les dimensions de temps et de fréquence s'expriment respectivement via la notion de rythme et de mélodie. Deux aspects sont concernés : la résolution locale et l'intégration globale des séquences auditives. La capacité locale de résolution temporelle du système auditif possède une vision "microscopique". Dans cette voie de recherche, Hirsh mesure les seuils d'identification de l'ordre de sons, sifflements et clics [187], Miller et Dexter discutent l'intelligibilité de la parole dans l'ordre de discrimination des consonnes [188], et Winckel observe une permutation de l'ordre lors d'une séquence musicale, exécutée à tempo rapide [189]. La deuxième approche, l'intégration globale (ou encore holistique) traite les sons globalement et cette capacité permet la ségrégation et la formation de flux auditifs. On a mentionné dans le chapitre sur la synthèse granulaire que l'écoute globale fonctionne quand le système auditif n'arrive pas à "décortiquer" en détail les composantes de la mixture sonore. L'intégration globale semble être plus utilisée par l'oreille que la résolution locale, car elle permet un traitement plus rapide de l'information sonore [190]. L'intégration globale est associée souvent à la perception "primitive", alors que la résolution locale demande un effort d'attention plus important.

11.1.4 Attention auditive

L'attention traduit l'acte de se concentrer sur un stimulus afin de mieux le percevoir et l'observer. Face aux plusieurs stimuli, notre comportement est soumis à deux types d'attention : l'attention sélective et l'attention partagée. L'attention sélective exige la concentration sur un seul stimulus, alors que l'attention partagée, illustre la compétition entre plusieurs stimuli. Les théories de l'attention associées à la perception des séquences auditives peuvent être regroupées dans trois approches [186].

Théories des étapes et Gestalt La première approche, établie par Bregman dans le but d'expliquer la séparation de sons en provenance de différentes sources, est une variante de la théorie des étapes, combinée avec les concepts de Gestalt. Cette approche considère deux phases. Lors de la première phase, dite pré-attentive, l'auditeur crée un regroupement des objets sonores selon les principes de Gestalt de proximité et de continuité. La deuxième phase est marquée par une activation des schémas pré-acquis qui guident l'attention. Bregman souligne l'importance d'un "temps d'attente" pour que l'auditeur s'adapte à la séquence écoutée et commence à percevoir la séparation des flux.

Relation temps-fréquence La deuxième approche, issue des études de Van Noorden sur les flux auditifs, considère que le temps et la fréquence forment perceptivement une seule unité d'un ordre supérieur : la relation temps-fréquence. Les changements fréquentiels reflètent sur les jugements temporels.

Attention dynamique La troisième approche, proposée par Jones, se focalise sur les aspects dynamiques de l'attention [191]. L'approche dynamique étudie l'évolution en temps des objets sonores. Elle suppose deux types de relations temporelles à l'intérieur des événements : de mouvement et de rythme. Les relations de type mouvement expriment l'évolution au cours du temps d'une dimension perceptive des

sons, par exemple la hauteur ou la sonie. L'attention associée au mouvement de la fréquence détermine trois régions fréquentielles, en accord avec les seuils de cohérence et fusion déterminés par Van Noorden. Une cadence fréquentielle petite permet le suivi de groupes successifs de tons répétés. La région intermédiaire correspond au concept d'un trille, où le mouvement de la fréquence peut encore être suivi. Les grandes variations en fréquence brisent la ligne de mouvement continu, empêchent l'attention synchrone et créent une représentation parallèle, i.e les flux auditifs. Le rythme apparaît conjointement au processus de mouvement. En effet, chaque changement de la fréquence, de l'intensité, du timbre ou de la durée des sons crée des ruptures lors de l'évolution temporelle des événements. Ces changements créent la sensation d'accents rythmiques à l'intérieur d'une séquence temporelle. Les accents sont utilisés par l'attention comme des repères pour les périodes en temps d'ordre supérieur, i.e plus globales. Autrement dit, ces accents permettent à l'auditeur de ne plus se concentrer finement sur chacun des sons dans un groupe, mais de passer à une écoute plus globale, en se basant sur la périodicité créée par les groupes rythmiques.

11.2 Le bruit de moteur dans le contexte des scènes auditives

L'interprétation des bruits moteur au ralenti dans le contexte des scènes auditives est une approche nouvelle. Afin de l'introduire, cette section relie les principes de Gestalt, la formation de flux auditifs, les notions d'intégration et de rythme avec les bruits moteur au ralenti.

11.2.1 Bruit moteur et flux auditifs

Le rapprochement entre les bruits moteur et les flux auditifs est l'objectif principal de notre étude perceptive. Trois expériences de Bregman sur la formation des flux auditifs vont illustrer cette idée [38].

Séquences de bruits La formation des flux auditifs s'effectue non seulement dans les séquences de sons, mais aussi dans les séquences de bruits. D'une manière analogue à l'exemple avec les sons, la séquence sonore ici contient successivement des bruits dans les aigus et les graves (**Fig. 11.5 a,a'**). De la même façon, à un tempo lent, l'oreille arrive à suivre l'enchaînement aigus-graves (**Fig. 11.5 a**), alors qu'à un tempo rapide, deux flux se forment (**Fig. 11.5 a'**). L'exemple sonore se trouve sur (**Piste 78**).

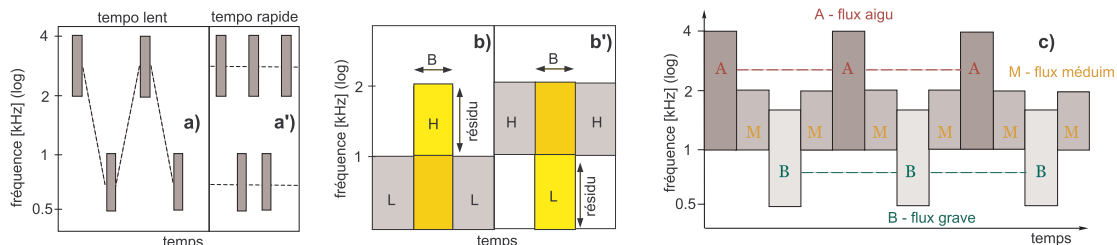


FIG. 11.5 – Flux auditifs avec des séquences de bruits bande étroites (d'après Bregman [38])

Formation de résidu Le stimulus principal dans cet exemple représente un bruit large bande B (**Fig. 11.5 b,b'**) qui peut être décomposé en deux sous-bandes, une bande aiguë H et une bande grave L, de manière que $B=L+H$. Dans la séquence, le bruit B est précédé et suivi soit par un bruit grave de type L (**Fig. 11.5 b**), soit par un bruit aigu de type H (**Fig. 11.5 b'**). Le résultat sonore de l'expérience (**Piste 79**) montre que la partie grave du bruit B est "attirée" par le bruit grave L, en faisant ressortir un résidu impulsif aigu H qui apparaît une fois pour chaque cycle (**Fig. 11.5 b**). Similairement, la bande aiguë H attire les fréquences aiguës du bruit B et forme un résidu impulsif dans les graves L (**Fig. 11.5 b'**).

Formation de trois flux La troisième expérience présentée ici utilise trois types de bruit de largeur de bande variable. Le premier bruit, de type A, couvre les fréquences aiguës et médiums (1-4 kHz). Le bruit de type M contient les médiums (1-2 kHz), alors que le bruit dans les graves B s'étend sur la bande (0.5-2 kHz) (**Fig. 11.5, c**). L'ordre de répétition de la séquence est AMBM .. (**Piste 80**) fait ressortir trois flux distincts : aigu (A), médium (M) et grave (B).

Lien avec les bruits moteur Ces exemples mettent en évidence le rapprochement entre les bruits moteur et la formation des flux auditifs. En effet, le cycle moteur représente une séquence d'événements sonores, où chaque événement possède un spectre spécifique, intervient à un moment précis dans le cycle

moteur et se répète d'une manière périodique. On peut imaginer le bruit moteur comme une séquence similaire à l'exemple de (**Fig. 11.5, c**), où les flux A-A-A ..., M-M-M ... ou B-B-B ... correspondent aux bruits issus d'une source sonore commune, par exemple les explosions, les claquements de soupapes. La répétition rapide de la séquence du ralenti moteur mène à la ségrégation des événements en plusieurs flux auditifs. A un régime moteur très élevé (> 3000 tr/min), l'oreille ne parvenant pas à séparer en temps les événements, perçoit les flux en continu comme des harmoniques du moteur (**Piste 81**).

11.2.2 Structure du bruit moteur

La structure temporelle du bruit moteur peut être décomposée en trois niveaux : l'explosion, le cycle moteur et le bruit moteur (**Fig. 11.6**). On peut supposer que la perception prend en compte ces niveaux.

Niveau 1 : un quart de cycle moteur Le quart du cycle moteur, dont la contribution majeure vient de l'explosion, représente un bruit impulsif d'une durée entre 40 ms au ralenti et moins de 10 ms pour un régime supérieur à 3000 tr/min. Le bruit moteur est une séquence de bruits impulsifs de très courte durée. Chaque explosion possède un contenu spectral et un profil temporel spécifiques déterminés par la loi de combustion dans le cylindre. Au ralenti, les explosions sont perceptibles par l'oreille et même une hauteur peut leur être attribuée, comme le montrent les exemples sonores, issu d'un bruit enregistré au ralenti, et représentant quatre quarts d'un cycle moteur **Piste 82, Piste 83, Piste 84 et Piste 85**. En revanche, l'oreille ne perçoit pas en détail les événements se produisant à l'intérieur de l'explosion. Le quart de cycle moteur peut être considéré donc comme la plus petite unité perceptible pour l'oreille.

Niveau 2 : cycle moteur Le cycle moteur représente la succession de 4 explosions différentes pour un moteur à 4 cylindres. La différence entre les explosions rend le bruit moteur moins régulier et impulsif par rapport au bruit d'un monocylindre. Les différences entre les niveaux sonores et les timbres des explosions créent un motif rythmique propre à une configuration pour chaque moteur.

Niveau 3 : bruit moteur Le troisième niveau introduit la répétition du cycle moteur. Cette répétition résulte un rythme spécifique du bruit, déterminé par les différences perceptives entre les cylindres. A ce niveau se produit la légère variation des explosions de chaque cylindre d'un cycle à l'autre. Le bruit global possède une sonie et timbre qui vont dépendre des sonies et les timbres des explosions.

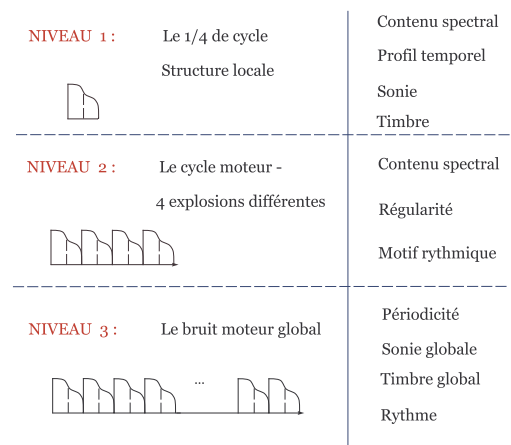


Fig. 11.6 – Structure temporelle du bruit moteur

11.2.3 Notion de rythme

La définition du rythme le représente comme la répétition périodique d'un événement. Dans le contexte de la musique, le rythme est déterminé par la successions de temps accentués et non accentués. Le rythme dans le contexte des bruits moteur va être déterminé comme la succession d'explosions possédant des sonies et timbres différents.

Comme on l'a déjà noté, l'approche dynamique de l'attention auditive utilise les accents comme des repères permettant à l'auditeur de regrouper les événements dans une structure cyclique. Une relation simple s'établit quand les événements sonores surviennent avec une régularité périodique, dite hiérarchique [186]. Les groupements réguliers possédant un rapport constant et entier entre une période supérieure et inférieure (en l'occurrence $T_n/T_{n-1}=4$) sont appelés encore harmoniques (**Fig. 11.7**).

Ce rythme composé de 4 temps est illustré à l'aide d'un rythme de samba (**Piste 86**) et comparé avec le bruit ralenti du ralenti diesel (**Piste 87**). La ressemblance est nette, toutefois, le rythme de samba est rythmiquement plus irrégulier que le rythme marqué par le ralenti diesel.

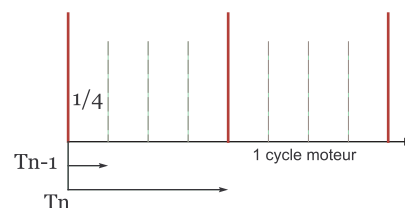


Fig. 11.7 – Relation rythmique du cycle moteur

11.3 Exemples rythmiques au ralenti

La méthode d'analyse-synthèse, présentée dans la deuxième partie de la thèse, permet la manipulation directe des atomes temps-fréquence. En s'inspirant des exemples sur les flux auditifs, notre objectif sera de montrer les effets sur la perception dus à la soustraction de certaines zones temps-fréquence dans les bruits de synthèse. La soustraction se fait en éliminant les atomes dont les "centres" se trouvent dans un rectangle temps-fréquence, déterminé par des limites en temps et en fréquence. L'intérêt d'une telle expérience est, d'abord, de pouvoir observer l'influence perceptive d'une zone temps-fréquence sur le bruit de synthèse final et, par la suite, de pouvoir associer des événements physiques temps-fréquence dans le cycle moteur à leur signature sonore et rythmique.

11.3.1 Filtrage coupe-bande

Cet exemple présente la soustraction de la bande fréquentielle de 2-5 kHz. A partir du son de synthèse (**Piste 88**) d'un cycle moteur (**Fig. 11.8 a**), où les astérisques blancs affichent les atomes retenus pour la synthèse, on soustrait les atomes dont les centres sont situés entre 2 et 5 kHz (**Fig. 11.8 b**).

Cette soustraction représente une sorte de filtrage coupe-bande (**Fig. 11.8 c**), puisqu'on agit localement sur le spectre d'un cycle. Le résultat sonore après la soustraction (**Piste 89**) montre que le timbre du bruit devient un peu plus feutré, toutefois l'aspect impulsif du bruit est conservé. Ce résultat est similaire sur un autre bruit de synthèse avant (**Piste 90**) et après (**Piste 91**) la soustraction de la bande 1-4 kHz. D'une manière générale, la qualité du bruit au ralenti ne dépend pas seulement du contenu spectral, mais également de l'aspect impulsif. Un filtrage coupe-bande peut améliorer perceptivement les bruits, car il enlève une partie très énergétique du spectre qui peut provoquer la gêne. Toutefois, une soustraction peut aussi laisser apparaître des phénomènes sonores contenus dans les hautes ou les basses fréquences, masquée auparavant par la sélection soustraite.

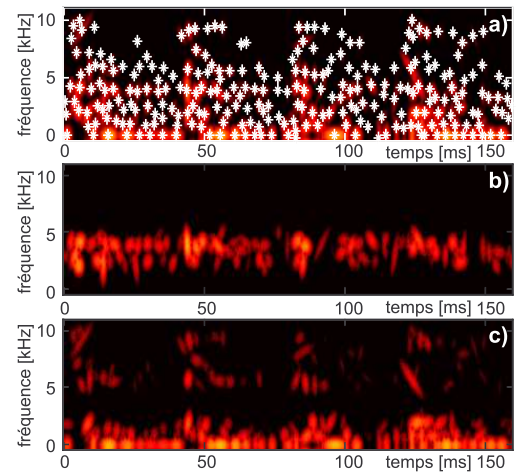


Fig. 11.8 – Bruit de synthèse (a), bande soustraite 2-5kHz (b), synthèse sans la bande (c)

11.3.2 Soustraction des atomes

La soustraction locale de quelques atomes peut influencer l'impulsivité du bruit. A partir de deux sons de synthèse (**Piste 88** et **Piste 92**), trois atomes sont soustraits, respectivement dans la zone [43-45 ms][1-7 kHz] (**Fig. 11.9 b**) et dans la zone [20-35 ms][2-4 kHz] (**Fig. 11.9 e**). Les bruits résultants (**Piste 93** et **Piste 94**) semblent plus impulsifs. La soustraction de 3 atomes, couvrant une portion temps-fréquence de quelques ms peut apporter des changements perceptifs.

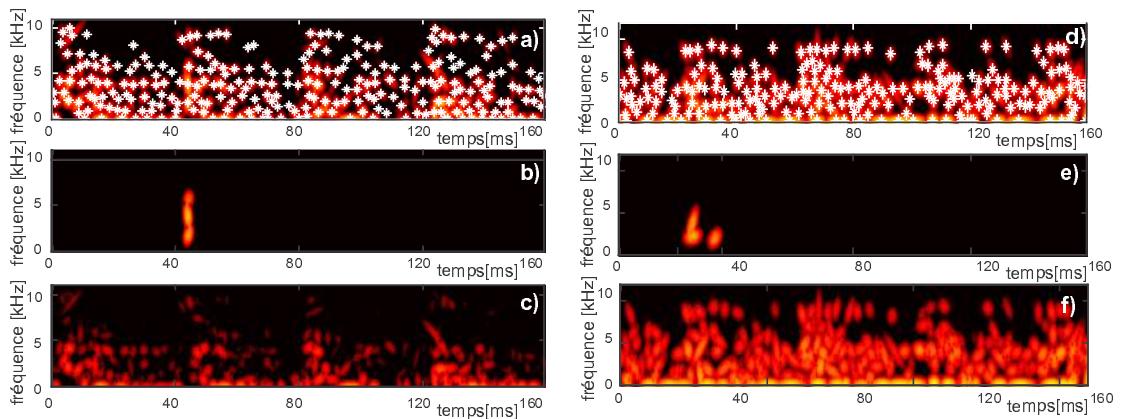


FIG. 11.9 – Signaux de synthèse avant (a,d) et après (c,f) la soustraction de 3 atomes dans les zones [43-45 ms][1-7 kHz] (b) et [20-35 ms][2-4 kHz] (e).

11.3.3 Soustraction d'une zone temps-fréquence

La soustraction d'une zone temps-fréquence plus grande en temps peut influencer davantage l'impulsivité du bruit. Deux facteurs interviennent dans ce processus : l'énergie spectrale contenue dans la zone temps-fréquence et l'endroit en temps où se produit la soustraction. Pour illustrer les effets perceptifs sur l'impulsivité et le rythme des bruits au ralenti, on va soustraire une zone de 20 ms en temps et on va utiliser deux largeurs de bande fréquentielle, [0-12 kHz] et [1-5 kHz]. Cette sélection va être éliminée aux différents endroits du cycle moteur, afin de montrer l'influence perceptive sur le rythme.

La zone [0-20 ms] et [0-12 kHz] est soustraite (**Fig. 11.10 b**) du signal de synthèse (**Fig. 11.10 a**). Le résultat de la soustraction (**Fig. 11.10 c**) témoigne d'une impulsivité beaucoup plus importante dans le bruit (**Piste 95**) par rapport au bruit de synthèse initial (**Piste 88**). Ce phénomène est très important, car il montre que le fait d'enlever de l'énergie du signal, peut augmenter l'impulsivité du bruit au ralenti et, par conséquent, le rendre plus désagréable. Ce désagrément peut être d'autant plus grand que l'énergie contenue dans la zone soustraite est importante.

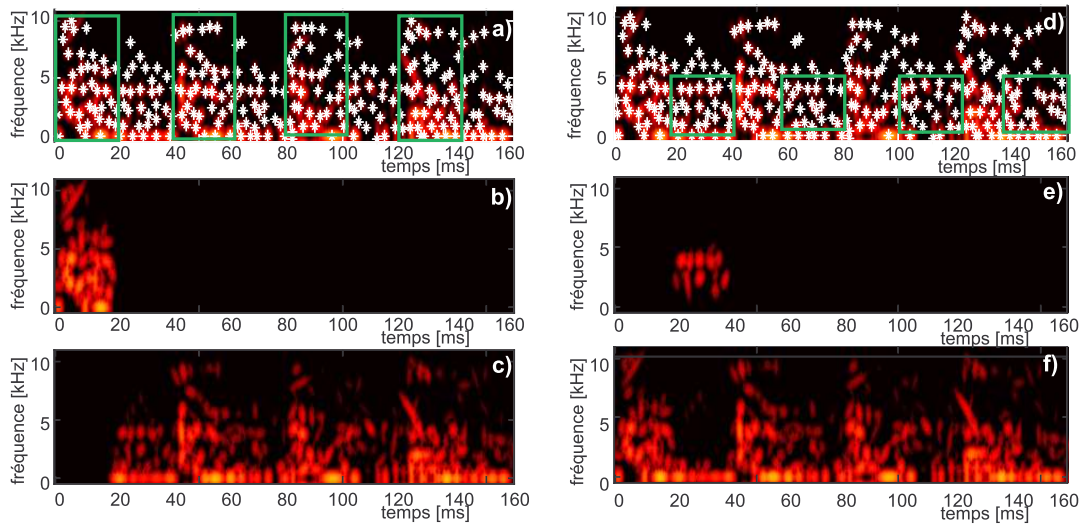


FIG. 11.10 – Soustraction de zone temps-fréquence. A gauche : signal de synthèse (a), zone sélectionnée [0-20 ms] [0-12 kHz] (b) et le résultat de la soustraction (c). Identique à droite pour la zone [20-40] [1-5 kHz].

Afin d'illustrer l'influence sur le rythme de la séquence, on va soustraire une zone de la même taille (20ms, 12 kHz) aux endroits différents du cycle moteur. Ces sélection sont marquées avec des rectangles verts sur la (**Fig. 11.10 a**). Les sons de synthèse avec une zone soustraite [40-60 ms][0-12 kHz], [80-100 ms][1-12 kHz] et [120-140 ms][1-12 kHz] se trouvent respectivement sur (**Piste 96**), (**Piste 97**) et (**Piste 98**). Ils témoignent d'une impulsivité augmentée et d'un rythme différent engendré suite à la soustraction. Le silence engendré par la soustraction sur 20 ms reste nettement perceptible et dans le bruit final, on a l'impression qu'il "saute" un temps. La sélection plus étroite en fréquence [0-20ms][1-5kHz] provoque un changement moins marqué (**Piste 99**).

Les exemples précédents éliminent une sélection au début d'une explosion moteur. La soustraction de la deuxième moitié de l'explosion peut avoir un effet perceptif différent. En sélectionnant une zone plus petite en fréquence, de (1-5 kHz) (**Fig. 11.10 f**) et en conservant la même durée de 20 ms en temps, la soustraction de [20-40 ms][1-5 kHz] n'apporte pas un changement perceptif important (**Piste 100**) par rapport au bruit de synthèse initial (**Piste 88**). Ceci est probablement dû à une énergie moins importante dans la zone sélectionnée, de la bande fréquentielle moins large, ainsi que d'un phénomène de masquage effectué par la zone très énergétique de (0-20) ms, qui précède notre sélection. De même, la soustraction des zones [60-80 ms][1-5 kHz], [100-120 ms][1-5 kHz] et [140-160 ms][1-5 kHz], marquées par des rectangles (**Fig. 11.10 d**) n'apporte pas de changement perceptif important (**Piste 101**, **Piste 102** et **Piste 103**).

En conclusion, cet exemple montre que la soustraction d'une zone temps-fréquence peut avoir une influence sur le rythme et l'impulsivité du bruit, en fonction de la largeur de la bande fréquentielle et l'endroit (début ou fin d'explosion) où elle est effectuée. Il semble qu'une soustraction sur toute la plage [0-12 kHz] est beaucoup trop "violente", car elle élimine également les fréquences en dessous de 1 kHz qui contribuent à une certaine "fluidité" du bruit.

11.3.4 Soustraction de plusieurs zones

La soustraction de plusieurs zones dépend également du contenu fréquentiel et l'endroit en temps où s'effectue la sélection. Cette modification semble ne pas avoir une influence importante sur le rythme si la soustraction se fait d'une manière équidistante en temps. La distance est constante entre les zones [10-30; 50-70; 90-110; 130-150 ms][1-5 kHz] et entre les zones [20-40; 60-80; 100-120; 140-160 ms][1-5 kHz], comme le montrent respectivement la (Fig. 11.11 b) et (Fig. 11.11 e). Les sons de synthèse après la soustraction, **Piste 104** et **Piste 105**, montrent une petite différence perceptive par rapport au bruit de synthèse initial (**Piste 88**). Ces différences concernent essentiellement le timbre et la hauteur des bruits.

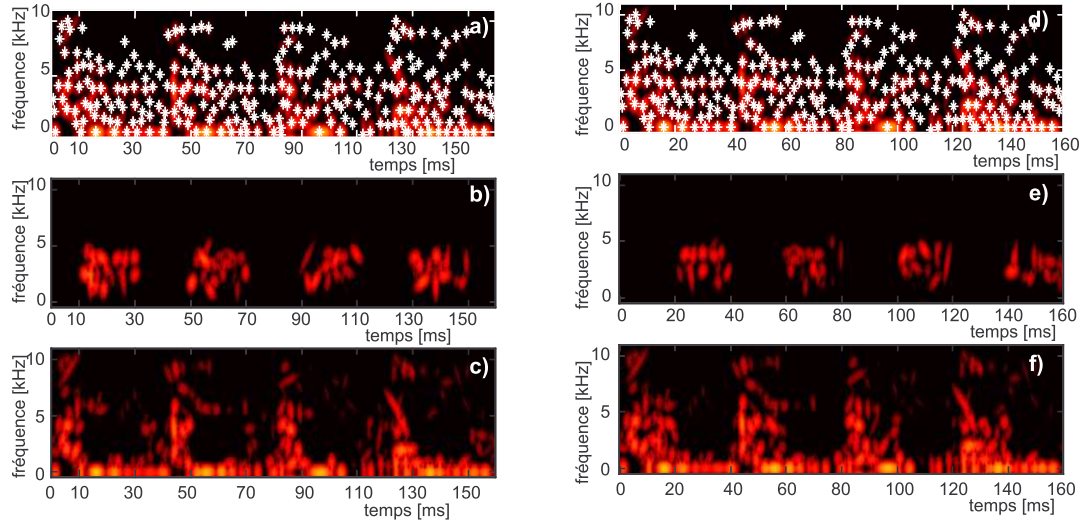


FIG. 11.11 – Soustraction de quatre zones temps-fréquence. A gauche : zones [10-30 ; 50-70 ; 90-110 ; 130-150 ms][1-5 kHz] . A droite : zones [20-40 ; 60-80 ; 100-120 ; 140-160 ms][1-5 kHz]

L'influence perceptive n'est pas très significative probablement parce que les zones soustraites sont disposées d'une manière équidistante en temps et parce que les deux soustractions conservent le début des explosions du moteur, en enlevant les atomes dans une partie où l'explosion possède moins d'énergie (Fig. 11.11 c,f).

Un exemple de soustraction alternative dans les basses et les hautes fréquences apporte un changement significatif dans la perception du rythme (Fig. 11.12). La soustraction au début de chaque explosions, alternant les zones de [0-10; 80-90 ms] dans les basses et moyennes fréquences [0-4 kHz] et [40-50; 120-130 ms] dans les hautes fréquences [4-12 kHz] provoque un effet de séparation en flux auditifs : un dans les basses fréquences et un dans les hautes fréquences (**Piste 106**). Ces flux semblent plus éloignés en fréquence que dans le bruit sans modification (**Piste 88**). De plus, le rythme semble plus irrégulier, plus impulsif où le flux dans les hautes fréquence est perçu encore plus nettement.

En conclusion, la soustraction d'une zone temps-fréquence apporte des changements perceptifs dans le timbre, la hauteur, l'impulsivité et le rythme du bruit au ralenti. Ces changements dépendent de l'endroit, de la taille et du contenu de la zone sélectionnée.

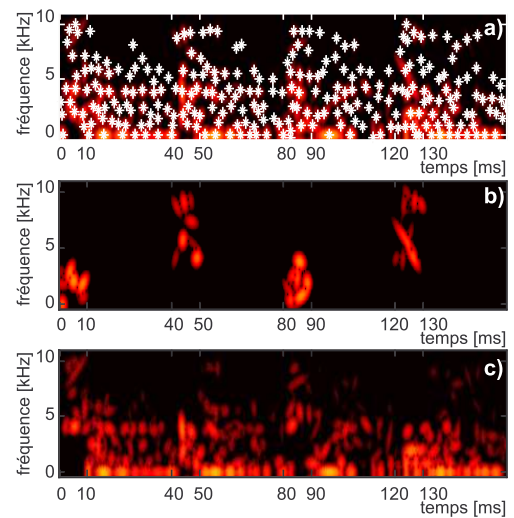


Fig. 11.12 – La soustraction alternative des zones et [40-50 ; 120-130 ms][4-12 kHz]

Conclusion sur les scènes auditives

Dans ce chapitre j'ai présenté la perception des séquences de sons dans le contexte des scènes auditives. D'après ce concept, les sons sont groupés par l'oreille selon des lois de Gestalt de proximité, de similarité, de continuité, de clôture et de "destin commun". La perception de séquences, constituées de sons graves et aigus présentés en alternance, met en évidence la formation de flux auditifs lors d'une exécution rapide de la séquence. Cette idée repose à la base de notre approche des bruits moteur au ralenti, puisque plusieurs exemples sonores permettent de faire le lien entre les flux auditifs et la structure des bruits au ralenti. La méthode d'analyse-synthèse développée dans la thèse a permis d'expérimenter avec la soustraction d'atomes dans les différentes zones temps-fréquence du signal de synthèse. Ces exemples permettent d'étudier l'influence perceptive sur le timbre, la hauteur, la sonie et le rythme des séquences de bruit au ralenti.

Orientation de la thèse : L'hypothèse de la formation des flux auditifs dans les basses, moyennes et les hautes fréquences des bruits moteur au ralenti est l'idée principale des tests perceptifs, présentés dans le chapitre suivant. Mon objectif sera de montrer l'influence des différentes zones temps-fréquence et la formation de flux sur la perception du timbre, la sonie et l'impulsivité des bruits aux ralenti. L'intérêt principal de cette approche réside dans le fait de pouvoir associer les phénomènes physiques dans le cycle moteur avec leur signature sonore dans une zone temps-fréquence. En dépassant le sujet de la présente thèse, l'objectif ambitieux sera de pouvoir prédire l'impact des modifications des phénomènes physiques dans le cycle moteur sur la perception des bruits. La théorie des flux auditifs peut aussi donner une nouvelle approche dans l'étude des bruits impulsifs répétitifs.

Tests perceptifs pour les bruits au ralenti

Ce chapitre représente l'étape finale de la thèse : l'application de la méthode d'analyse-synthèse temps-fréquence pour l'étude perceptive des bruits au ralenti diesel. La perception des bruits de synthèse est étudiée à l'aide de deux types de tests : test de catégorisation et test de préférence. Les bruits de synthèse pour les tests sont construits en modifiant certaines zones temps-fréquence. Les modifications visent à tester l'impact sur la perception des zones en fonction de leur position en temps dans le cycle moteur, leur contenu spectral, leur niveau sonore et leur interaction avec les autres zones. L'objectif des tests est de trouver des indicateurs perceptifs, permettant d'expliquer la perception des bruits au ralenti, ainsi que de mettre en évidence la prise en compte des flux auditifs et le rythme dans le jugement perceptif. Le chapitre présente consécutivement : le découpage schématique du bruit moteur en zones temps-fréquence, la construction des sons pour les tests, les outils statistiques utilisés et enfin, les résultats des tests auditifs.

12.1 Cycle moteur : étude en quarts

Le concept de zones temps-fréquence est introduit afin de schématiser géométriquement le bruit moteur sur une représentation temps-fréquence. Il est effectué en deux étapes : en découpant le cycle moteur en quatre parties en temps et en découpant chaque quart du cycle en 4 zones temps-fréquence.

12.1.1 Découpage du cycle en quarts

Le cycle moteur est une succession de 4 sons impulsifs, générés par les explosions. En découpant temporellement le cycle en quatre parties égales, on associe chaque quart de cycle moteur à une des explosions dans les cylindres. Toutefois, un quart de cycle moteur contient non seulement les explosions, mais aussi les autres phénomènes physiques survenant dans le cycle moteur comme l'injection et le bruit des soupapes. Notre démarche ne cherchera pas à séparer la contribution de ces différentes sources, mais de regrouper les événements dans le cycle moteur par portions en temps bien perceptibles par l'oreille.

La durée du cycle moteur pour un régime en ralenti de 750 tr/min, est de 160 ms en moyenne. Les quarts de cycle s'étalent donc sur environ 40 ms. Des sons de cette durée sont bien perceptible par l'oreille qui peut leur attribuer également un timbre, une hauteur et un niveau spécifiques. Chaque explosion se traduit par un bruit impulsif, composé d'une partie d'impact et d'une partie d'amortissement, qu'on va noter respectivement "attaque" et "amortie". Par simplicité, on va considérer que chacune de ces deux parties occupe en temps la moitié du quart de cycle moteur, i.e environ 20 ms pour un bruit au ralenti (**Fig. 12.1**). La partie "attaque" représente une excitation large bande possédant une grande énergie et dont la contribution majeure se situe spectralement entre 1 et 5 kHz. En comparaison, la partie "amortie" possède très peu d'énergie et son spectre n'est pas aussi riche dans les hautes fréquences.

D'une manière générale, l'impulsivité et le rythme du bruit moteur dépendront des différences de niveau et de timbre entre les quarts de cycle moteur, mais également entre leurs parties "attaque" et "amortie" respectives. Le chapitre précédent a montré que les différences entre les niveaux des explosions vont déterminer le motif rythmique de la séquence du bruit moteur, puisque le rythme est défini comme la succession de temps plus et moins accentués. Dans l'objectif d'observer les différences de niveau sonore cycle-à-cycle et cylindre-à-cylindre, on va étudier les sonies spécifiques des quarts de cycle.

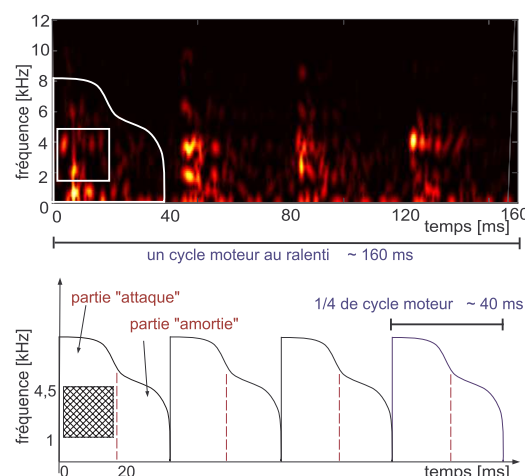


Fig. 12.1 – Découpage du spectrogramme d'un cycle moteur au ralenti en quarts

12.1.2 Sonie spécifique des quarts

La sonie spécifique des quarts de cycle moteur montrent les différences entre les explosions de chaque cylindre, ainsi que la signature caractéristique des cylindres d'un cycle à l'autre (**Fig. 12.2**). Cet exemple utilise le découpage en quarts de 4 cycles moteur consécutifs, extraits à partir d'un bruit moteur au ralenti enregistré. L'indicateur psychoacoustique de sonie spécifique montre, selon le modèle de Zwicker, la répartition de la sonie en fonction des 24 bandes de Barks. Les ressemblances cycle-à-cycle des quarts (verticalement) témoignent que chaque cylindre a sa propre signature. Les différences quart-à-quart (horizontalement) montrent que chaque cylindre est différent.

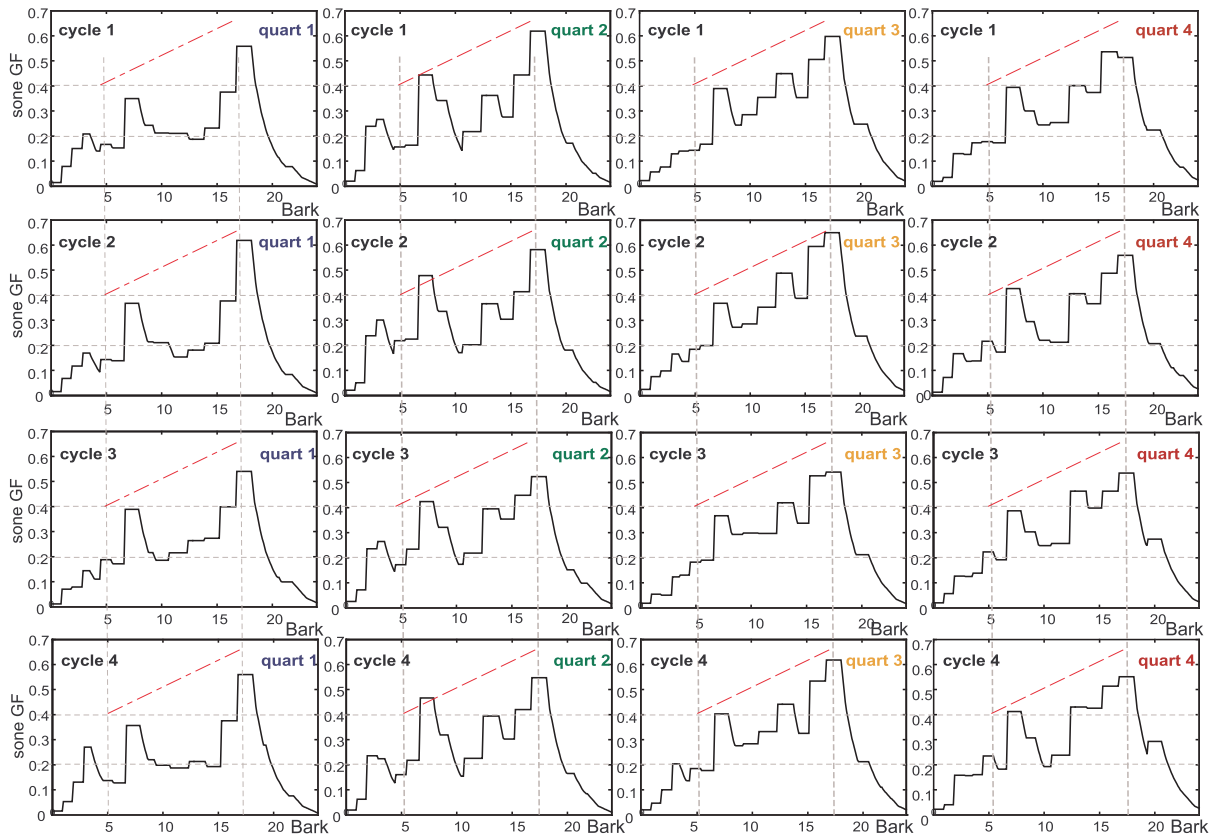


FIG. 12.2 – Sonie spécifique des quarts de cycle moteur sur quatre cycles consécutifs d'un bruit moteur diesel au ralenti. Verticalement : différences cycle-à-cycle ; horizontalement : différences cylindre-à-cylindre

Les différences entre la sonie des quarts s'expriment dans la répartition de la sonie spécifique dans l'aire contenu entre 7-ème et la 18 bandes de Barks. Les fréquences centrales correspondantes à ces bandes sont 700 et 4000 Hz. De plus, pour chaque quart de cycle, la sonie de la région autour de 4000 Hz est beaucoup plus importante que celle autour de 700 Hz. La répartition spectrale dans l'aire [4-18 Barks][0.2-0.7 Sones] est différente pour chaque quart de cycle. Elle semble être liée au centre de gravité spectrale de la sonie spécifique, ou encore le paramètre d'acuité. Les quarts de cycle possèdent donc une répartition spectrale qui déterminera leur timbre et leur hauteur spécifiques. Dans cet exemple, le quart 3 semble avoir la sonie la plus importante, alors que le quart 1, la moins importante. Cet exemple confirme l'importance de la plage fréquentielle de 1-4 kHz dans la perception du bruit moteur au ralenti. On va tenir compte de ces limites en fréquence pour le découpage spectral des quarts de cycle.

12.2 Concept de zones temps-fréquence

Afin de tenir compte de la répartition spectrale du bruit moteur, on va combiner le découpage temporel du cycle moteur en quarts avec un découpage du spectre du cycle moteur. Cette démarche nous permettra de définir un découpage en zones temps-fréquence du bruit moteur qui sera à la base de la construction des sons pour les tests perceptifs.

12.2.1 Définition

Le spectrogramme d'un quart du cycle moteur est découpé géométriquement en 4 zones temps-fréquence (**Fig. 12.3**). Dans ce découpage, la partie du spectre du bruit moteur située au-dessous de 1 kHz n'est pas prise en compte. Comme l'on montré les exemples du chapitre précédent, le fait d'enlever la plage fréquentielle entre [0-1 kHz] rend le bruit au ralenti très impulsif et moins naturel. La conservation de cette région permet de garder un fond continu dans les basses fréquences, propre au fonctionnement du moteur. De plus, ceci permet de concentrer l'attention sur les changements perceptifs plus importants se situant dans les moyennes et les hautes fréquences.

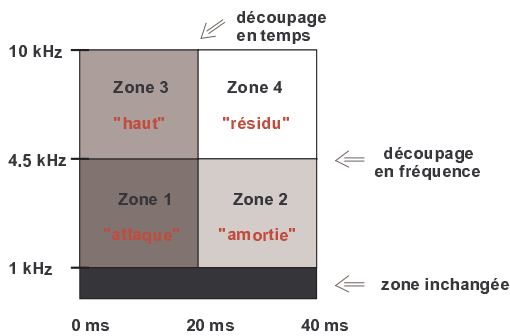


Fig. 12.3 – Découpage en 4 zones temps-fréquence du spectrogramme d'un quart du cycle

Les zones 1 et 2, "attaque" et "amortie", occupent temporellement la même région que dans le découpage précédemment défini des quarts de cycle, i.e respectivement de [0-20 ms] et [20-40 ms] du quart de cycle. Spectralement, ces deux zones contiennent les fréquences de [1-4,5 kHz]. Le choix des limites fréquentielles est déterminé par le fait que la zone 1-5 kHz contient l'énergie principale de la combustion. La limite choisie de 4,5 kHz suppose refléter un endroit "critique" pour la perception des bruits au ralenti. La limite supérieure, fixée à 10 kHz, permet de tenir compte des hautes fréquences entre [5-10 kHz], pouvant influencer grandement la perception, même si dans la pratique industrielle, les études de la combustion s'arrêtent même bien au-dessous, à environ 5-6 kHz. Au-delà de 10 kHz ne reste que très peu d'énergie significative dans le signal.

Les zones 3 et 4 sont situées respectivement dans le même créneau en temps que les zones 1 et 2, et comportent les hautes fréquences de [4,5-10 kHz] (**Fig. 12.3**). Le découpage est effectué d'une manière identique pour tous les quarts du cycle moteur et décompose le cycle moteur en $4 \times 4 = 16$ zones temps-fréquence au total (**Fig. 12.4**). Ce découpage schématique permet de garder un lien avec les phénomènes physiques dans le cycle moteur et avec la structure temps-fréquence des sons impulsifs. Cette approche permet aussi de regrouper les événements temporels de très courte durée qui se passent à l'intérieur de chaque zone et qui ne peuvent pas être perçus en détails par l'oreille. Le maillage temporel attribuant une durée de 20 ms pour chaque zone semble être approprié pour transmettre information du cycle moteur, car il reste dans la capacité de résolution temporelle de l'oreille.

Le découpage en zones temps-fréquence du cycle moteur permet de considérer le bruit moteur comme une séquence de sons, chacun occupant la région temps-fréquence d'une zone. Dans l'esprit des flux auditifs, on peut considérer que les zones qui ont un contenu spectral ou temporel proche, vont se regrouper et influencer la perception des bruits moteur au ralenti. Cette idée sera exploitée dans la construction des bruits de synthèse pour les tests perceptifs.

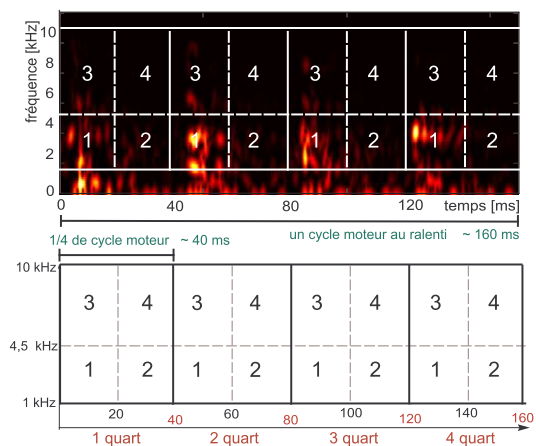


Fig. 12.4 – Découpage en zones temps-fréquence d'un cycle moteur au ralenti

12.2.2 Description des zones

Avant de passer à l'étape de construction des bruits de synthèse, on fera une description de chaque zone qui soulignera son importance, sa signification physique et sa contribution dans le niveau sonore et le rythme du bruit au ralenti.

Zone 1 : "attaque" C'est la zone la plus importante, car elle concentre la plus grande énergie du signal, due à la combustion. La zone "attaque" commence au début du quart de cycle moteur, et occupe le rectangle de [0-20 ms] et [1-4,5 kHz]. La région fréquentielle 1-4 kHz coïncide, comme le montrent les courbes d'isotonie de Fletcher et Munson, avec la zone la plus sensible de l'oreille [46]. Pour cette raison, la zone "attaque" peut avoir une contribution importante dans la gêne sonore. Grâce à son énergie importante, cette zone peut aussi effectuer un masquage de la zone la suivant temporellement, l'"amortie". Par ailleurs, l'impulsivité du bruit au ralenti est souvent liée à des différences de sonie et de timbre entre ces deux zones se succédant en temps. Dans ce sens, des différences de niveau et de timbre entre les zones "attaque" et "amortie" auront une grande influence sur l'aspect impulsif du bruit. Globalement, l'impulsivité sera d'autant plus importante que les zones "attaque" seront plus fortes par rapport aux autres zones.

Zone 2 : "amortie" Cette zone représente l'amortissement du son impulsif, dû à l'explosion. La zone "amortie" occupe la région [20-40 ms] et [1-4,5 kHz] et suit temporellement la zone "attaque" dans la même plage fréquentielle. D'une manière générale, l'énergie contenue dans la zone "amortie" est moins importante que celle de la zone "attaque". Cependant, la zone "amortie" est la deuxième zone la plus énergétique dans le quart de cycle moteur.

La différence de niveau et de timbre entre les zones "attaque" et "amortie" peut influencer l'appréciation du bruit au ralenti. Un niveau et un timbre relativement similaires entre ces deux zones peut altérer la perception de la périodicité globale, en rendant le bruit perçu deux fois plus rapide. Une grande différence de niveau entre "attaque" et "amortie" contribue à l'impulsivité du bruit. Ce phénomène apparaît clairement dans la comparaison entre des bruits au ralenti avec ou sans injection pilote (**Fig.12.5**). Le bruit sans injection pilote possède une combustion très violente qui se manifeste par une grande différence d'énergie entre les zones "attaque" et "amortie", où "attaque" contient presque toute l'énergie de l'explosion. Au contraire, la combustion avec une injection pilote est plus douce, ce qui se manifeste par une répartition équilibrée de l'énergie du signal entre les zones "attaque" et "amortie". Cette interaction entre les zones "attaque" et "amortie" semble très importante dans l'étude perceptive des bruits au ralenti.

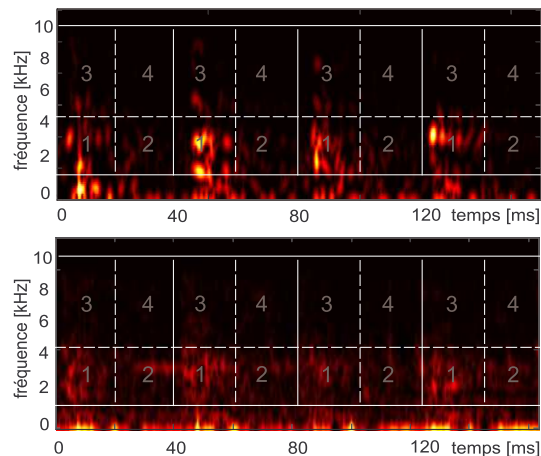


Fig. 12.5 – Deux cycles diesel au ralenti avec différents réglages de la combustion. Sans injection pilote (en haut), avec injection pilote (en bas)

Zone 3 : "haut" La zone "haut" est complémentaire à la zone "attaque", au sens où elle occupe le même créneau temporel [0-20 ms]. La zone "haut" contient les hautes fréquences de [4,5-10 kHz] qui sont généralement assez gênantes pour l'oreille. Elle peut contribuer à la perception du désagrément du bruit moteur au ralenti. Cette zone peut être aussi associée à l'impulsivité du bruit, car plus l'impact de l'explosion est fort, plus il sollicite la structure et fait apparaître des hautes fréquences. La présence des hautes fréquences dépendra des réglages et de l'architecture du moteur.

Zone 4 : "résidu" Cette zone possède souvent très peu ou pas du tout d'énergie du signal. Toutefois, elle peut être présente sur certains bruits diesel avec injection pilote. Il est intéressant d'étudier l'influence perceptive des bruits en présence de cette zone. Le fait que les bruits avec injection pilote soient assez bien perçus, laisse supposer que probablement la zone "résidu" effectue un "équilibre" temporel et fréquentiel du bruit. En effet, le "résidu" est une zone en décalage temporel et spectral avec la zone la plus énergétique, l'"attaque", et on peut imaginer, par association avec les scènes auditives de Bregman, que la présence plus importante de "résidu" et l'alternance des deux zones "attaque" et "résidu" favorisera la formation de flux auditifs. Cette question sera abordée plus loin.

12.2.3 Construction des sons

Le découpage en zones temps-fréquence du cycle moteur va donner, en combinaison avec la synthèse granulaire, une nouvelle manière d'étudier les sons impulsifs répétitifs. Dans le but de construire un ensemble de bruits de synthèse pour les tests perceptifs, notre démarche consiste à prendre un bruit de synthèse du ralenti moteur et d'effectuer des modifications dans certaines zones temps-fréquence. Ces modifications permettront d'étudier l'impact sur la perception des zones.

Choix de zones Le choix des zones à modifier est une étape importante, car toute zone peut avoir à priori, une influence sur la perception du bruit. La modification d'une seule zone permet d'associer un changement dans les paramètres du signal à un effet perceptif. La modification simultanée de plusieurs zones peut apporter plus d'informations sur les mécanismes d'interaction temporelle ou spectrale entre des zones et des phénomènes de masquage temps-fréquence. Toutefois, la modification simultanée de plus de deux zones devient très complexe, car il devient difficile d'expliquer l'influence des facteurs du signal sur les réponses perceptives. J'ai choisi de modifier séparément une zone temps-fréquence à la fois et d'étudier une interaction, entre les zones "attaque" et "amortie".

Choix de nombre de zones Le découpage en zones temps-fréquence donne 16 zones, associé aux quatre quarts de cycle moteur. Un cycle moteur comporte donc 4 zones "attaque", 4 zones "amortie", 4 zones "haut" et 4 zones "résidu". La modification de la zone "attaque" peut être faite sur un seul ou sur plusieurs cylindres. Les modifications peuvent donner une information sur l'effet perceptif des différences cylindre-à-cylindre. Cette approche introduit un grand nombre de combinaisons possibles associées aux modifications de chaque zone et de chaque cylindre. Pour des raisons de simplicité, on a décidé d'apporter la même modification dans une zone simultanément pour les 4 quarts du cycle moteur. Cependant, la différence cylindre-à-cylindre représente l'un des objets les plus intéressants pour une future étude.

Type de modification Plusieurs types de modifications peuvent être apportées sur les paramètres des atomes dans la zone temps-fréquence. Comme on l'a déjà vu, chaque atome est déterminé par 5 paramètres, pouvant être modifiés séparément : l'apparition en temps t_i , la fréquence ω_i , l'étendue σ_i , le paramètre de chirp β_i et l'amplitude a_i . La modification du temps et de la fréquence permettent d'étudier des effets de décalage temporel ou une modulation fréquentielle des événements. Un changement de l'étendue des atomes peut apporter une connaissance sur les phénomènes perceptifs dus au recouvrement des atomes dans le plan temps-fréquence et le support des atomes le plus adapté pour la synthèse granulaire du cycle moteur. Toutefois, la modification de l'amplitude des atomes semble être la meilleure approche pour étudier l'influence des zones sur le bruit, car ce sont les différences de niveau qui influencent le plus souvent l'impulsivité et le rythme du bruit. La modification consiste à amplifier ou diminuer l'amplitude des atomes d'une zone. Le niveau en dB d'amplification fait l'objet de quelques études préliminaires sur le niveau relatif des zones, la discrimination perceptive des modifications et l'égalisation en niveau de deux types de sons de synthèse. Ces études sont présentées par la suite.

Sélection des atomes La modification d'une zone temps-fréquence suppose une première étape de sélection des atomes situés dans cette zone. Le critère de sélection choisit les atomes dont les coordonnées en temps et fréquence des "centres" sont situées dans les limites temporelles et fréquentielles de la zone. La sélection des atomes de la zone "attaque" du premier quart du cycle est montré sur (Fig. 12.6). Les atomes des zones "attaque" des 4 cylindres sont sélectionnés à condition que leur temps d'apparition t_i soit dans les intervalles [0-20 ms], [40-60 ms], [80-100 ms] et [120-160 ms]. Pour les atomes ainsi choisis, la fréquence ω_i doit être comprise entre [1 - 4.5 kHz]. La condition sur les coordonnées des centres ne tient pas compte de l'étendue des atomes et le support de certains atomes peut se retrouver légèrement hors de la zone choisie. Une autre conséquence est l'addition d'une petite portion de l'énergie des atomes situés au début et à la fin du cycle lors de la synthèse, due à une répétition en boucle du cycle de synthèse.

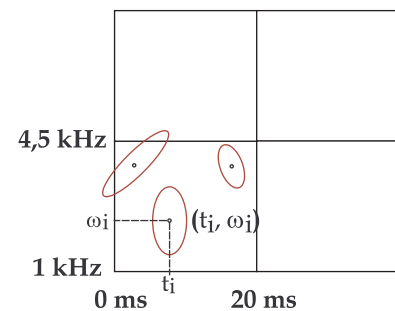


Fig. 12.6 – Sélection des atomes : la zone "attaque" du premier quart de cycle

Cependant, on peut considérer une telle sélection non rigoureuse des atomes sans conséquence importante sur la perception des sons de synthèse, car les atomes sont de courte durée et l'enveloppe gaussienne des atomes concentre plus l'énergie autour du "centre" de l'atome qu'aux extrémités du support.

Facteur de modification La modification d'amplitude des atomes est effectuée de la manière que chaque amplitude a_i de l'atome i est multipliée par un facteur qui correspond au niveau voulu d'amplification ou atténuation en dB. Le tableau récapitulatif des modifications, en fonction de la relation $(dB) = 20 \cdot \log(a)$ entre l'amplitude et l'échelle de dB se résume comme suit

Facteur d'amplitude	5,5	4	2	1	0,5	0.25	0.18
Modification en dB	+15 dB	12 dB	6 dB	0 dB	-6 dB	-12 dB	-15 dB

FIG. 12.7 – Modification d'amplitude des atomes : rapport entre le facteur de et le niveau en dB

Cette modification est relative à l'atome même. L'amplification ou l'atténuation totale du son de synthèse modifié va dépendre de l'énergie initiale contenue dans les atomes de la zone modifiée. L'amplification avec le même facteur va être plus importante pour une zone possédant au départ une énergie importante, que pour une zone de faible énergie. Le facteur d'amplification n'est pas le même pour les différents types de zone ("attaque", "amortie", etc), en revanche il est le même pour les 4 quarts du cycle.

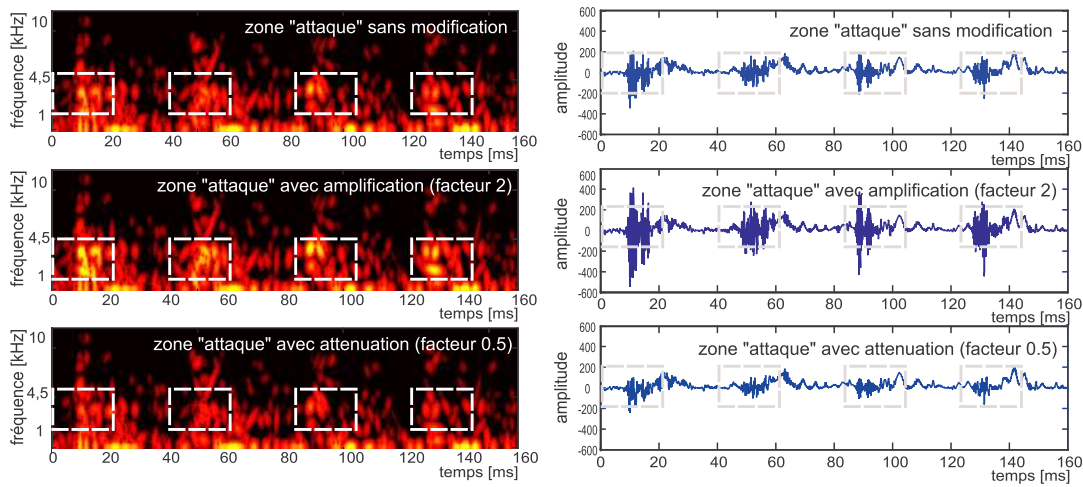


FIG. 12.8 – Modification de la zone "attaque" : sans modification (en haut), amplification à facteur 2 de l'amplitude des atomes (au milieu), atténuation à facteur 0.5 (en bas).

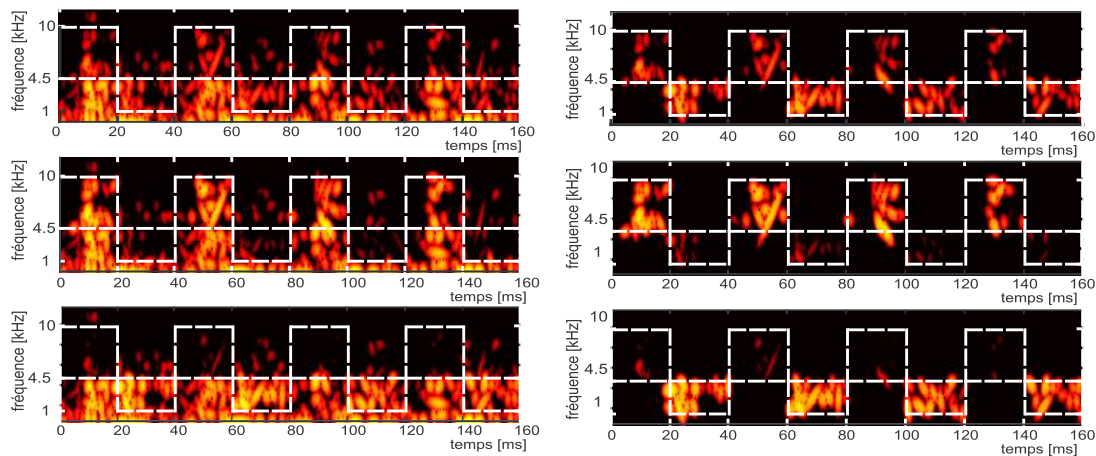


FIG. 12.9 – Amplification et atténuation alternative des zones "haut" et "amortie". A gauche, du haut en bas : le cycle moteur de synthèse; le cycle avec amplification de l'amplitude des atomes dans la zone "haut" et atténuation de la zone "amortie" de 15 dB; le cycle avec atténuation de l'amplitude des atomes dans la zone "haut" et amplification de la zone "amortie" de 15 dB. A droite : les zones modifiées

Les exemples d'amplification et atténuation de la zone "attaque" d'un facteur 2 (**Fig. 12.8**) et d'amplification et d'atténuation alternative des zones "haut" et "amortie" d'un facteur 5,5 (**Fig. 12.9**) illustrent l'idée de modification des zones temps-fréquence.

12.2.4 Niveau relatif des zones

La modification de l'amplitude des atomes dépend fortement de l'énergie du signal contenue dans la zone modifiée. Il est donc nécessaire de connaître le niveau sonore relatif des zones temps-fréquence. D'une manière générale, la zone la plus énergétique est l'"attaque", la zone la plus faible, le "résidu".

Signaux de synthèse Pour analyser les différences de niveau, on va utiliser deux signaux de synthèse, obtenus à partir de deux signaux originaux de bruit diesel ralenti avec deux réglages différents : le son 1 est sans injection pilote (**Piste 107**), le son 2 est avec injection pilote (**Piste 109**). Ces deux sons de synthèse (respectivement sur (**Piste 108** et **Piste 110**)) vont servir plus tard pour obtenir les deux familles de sons pour les tests perceptifs, appelés sons de "type 1" et sons de "type 2". L'intérêt d'observer les différences entre ces deux bruits provient du fait qu'il sont jugés perceptivement (par un groupe d'experts moteur) d'une manière opposée : le son 1 est jugé très désagréable, alors que le son 2 est jugé de bonne qualité.

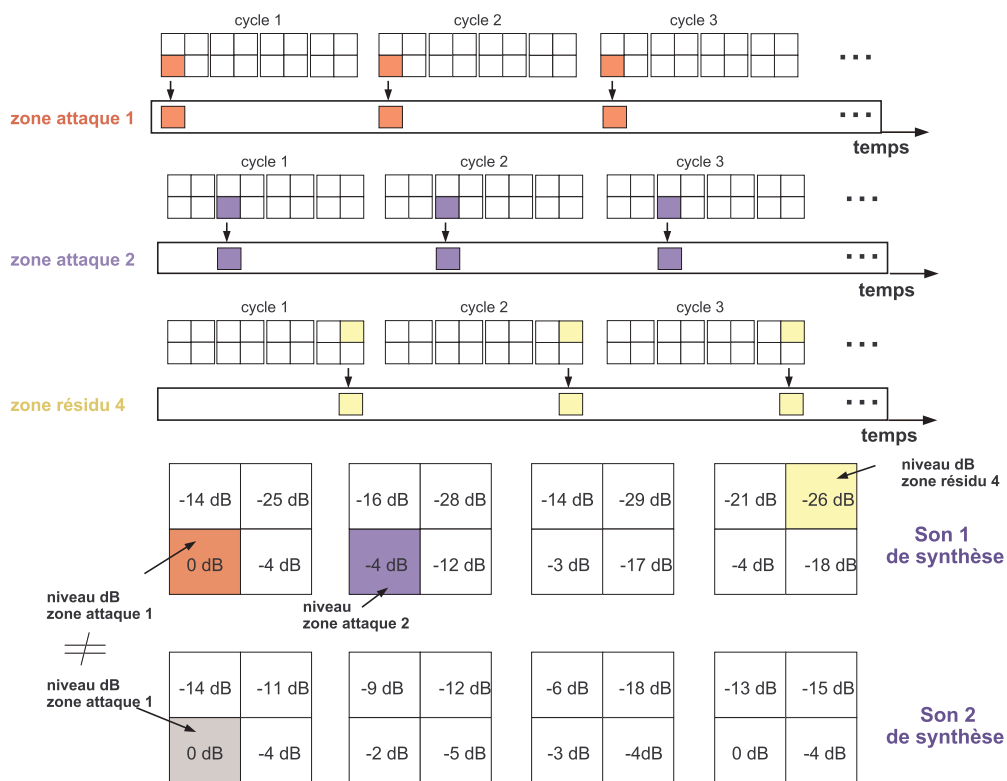


FIG. 12.10 – Calcul des niveaux relatifs à chaque zone temps-fréquence

Calcul des niveaux Le niveau sonore relatif à une zone, pour un quart de cycle, est estimé à la base des sons de synthèse contenant seulement cette zone. Le son relatif à la zone "attaque 1" est obtenu en retenant les atomes contenus seulement dans la zone "attaque" du premier quart de cycle; le son relatif à la zone "attaque 2" est obtenu en retenant les atomes contenus seulement dans la zone "attaque" du deuxième quart de cycle et ainsi de suite (**Fig. 12.10**). Au total, 16 sons de synthèse sont obtenus pour chacun des sons 1 et 2 (**Piste 111-142**). Pour chaque son 1 et 2, La zone possédant le niveau le plus élevée est prise comme un repère. En l'occurrence, la zone "attaque 1" (i.e la zone "attaque" du premier cylindre) est la plus forte pour les deux sons. On fixe à 0 dB, indépendamment pour le son 1 et le son 2, le niveau de leur zone "attaque 1". Par la suite, on garde le rapport le niveau relatif de tous les autres sons par rapport à ce niveau fixé de 0 dB. On compare les niveaux des zones seulement au sein du même son. Il faut préciser que le niveaux absolus de la zone "attaque 1" du son 1 et "attaque 1" de son 2 ne sont pas égaux (**Fig. 12.10**). Par ailleurs, le niveau global de son 1 est environ 3-4 dB plus élevée que le son 2, ce qui explique en partie l'appréciation du son 2.

Observations On peut constater que la zone "attaque 1" (0 dB), associée au cylindre 1 du son 1, est nettement plus forte les zones "attaque" des autres cylindres (-4 dB, -3 dB et -4 dB). Cette différence semble moins bien exprimée pour le son 2. On peut supposer que le son 1 possède un cylindre 1 qui est prépondérant sur les autres, alors que le son 2 possède deux cylindres forts, le premier et le quatrième. Cette différence de niveau peut avoir une influence sur le motif rythmique de chaque séquence. De plus, la différence de niveau entre la zone "amortie" et "attaque" est beaucoup plus grande pour le son 1 que pour le son 2. Cette observation peut expliquer l'impulsivité plus forte du son 1. La différence de niveau entre la zone "résidu" et "attaque" est deux fois plus importante pour le son 1 (~ 30 dB) que pour le son 2 (~ 15 dB). Il semble que l'énergie de son 1 est concentrée surtout dans la partie "attaque", alors que les autres zones restent plus faibles. Dans le son 2, l'énergie entre les zones est répartie d'une façon plus homogène. En conclusion, la répartition homogène de l'énergie semble favoriser l'appréciation du bruit, alors que les différences importantes de niveau entre les zones privilégient l'impulsivité.

12.2.5 Effet des zones sur le niveau global

La modification de l'amplitude des atomes d'une zone entraîne un changement du niveau global du son de synthèse. L'atténuation ou l'amplification de la zone a été défini en termes de multiplication de l'amplitude de chaque atome avec un facteur (**Fig. 12.7**). Toutefois, le fait d'augmenter de 6 dB l'amplitude des atomes d'une zone, n'augmente pas de 6 dB le niveau global du son de synthèse. De plus, le changement du niveau global est plus important lors de la modification d'une zone à forte énergie que d'une zone à faible énergie. L'étude de l'effet de modification de zone sur le niveau global utilise les mêmes sons de synthèse, son 1 et son 2. Cette étude a deux objectifs : déterminer les facteurs appropriés de modification pour chaque zone et assurer que les modifications apportent un effet perçu par les auditeurs.

Niveau relatif - niveau global La modification relative à l'amplitude des atomes apporte un changement dans le niveau global des sons de synthèse. L'amplification ou l'atténuation d'une seule zone à la fois aboutit, pour les 4 zones temps-fréquence, à la synthèse de 8 sons. Les modifications effectuées sur les atomes de chaque zone de son 1 et de son 2 se résument sur le tableau suivant

zone	amplification	atténuation
Attaque	+ 6 dB	-12 dB
Amortie	+12 dB	-6 dB
Haut	+6 dB	-6 dB
Residu	+6 dB	-6 dB

FIG. 12.11 – Facteur de modification de l'amplitude des atomes

Le niveau global en dB (en bleu) et dB(A) (en rouge) des sons de synthèse avec modification et des sons sans changement, est présenté sur (**Fig. 12.12**). Sur l'abscisse sont présentés d'abord les sons sans modification, et ensuite les sons avec modification respective des quatre zones : "attaque", "amortie", "haut" et "résidu". Le trait plain représente les amplifications, le hachuré -les atténuations.

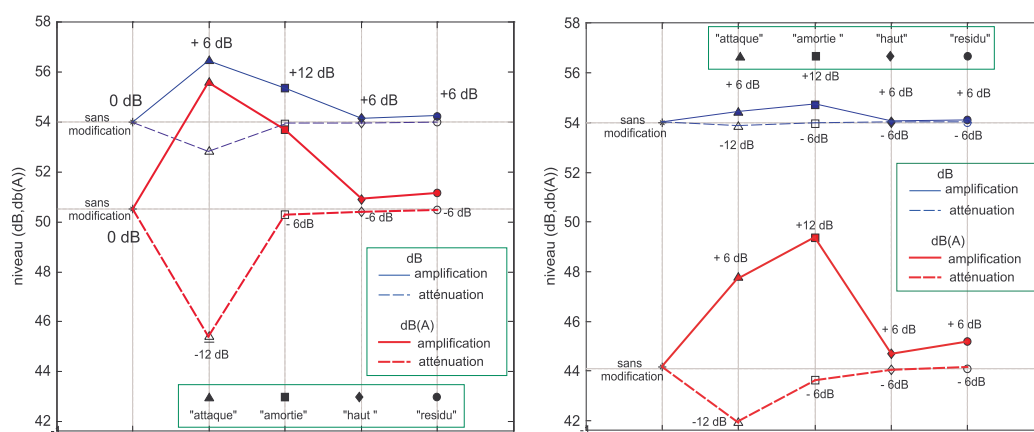


FIG. 12.12 – Niveaux des sons de synthèse après modification : son 1 (à gauche), son 2 (à droite)

Observations Les deux sons de synthèse sans changement sont égalisés à 54 dB. Toutefois, leurs niveaux en dB(A), de 50 dB(A) pour le son 1, et de 44 dB(A) pour le son 2, montrent que le son 1 possède une énergie plus grande. Par ailleurs, l'amplification plus importante de la zone "attaque" pour le son 1 par rapport au son 2 "attaque" est due à l'énergie plus importante contenue au départ dans cette zone. En revanche, l'amplification de la zone "amortie" a plus d'influence sur le niveau de son 2 que sur le son 1, car la zone amortie a été plus forte pour le son 2 que pour le son 1. Les modifications de ± 6 dB des zones "haut" et "résidu" n'influencent pas beaucoup le niveau en dB/dBA. Les sons de synthèse modifiés ne sont pas discriminables perceptivement ce qui nécessite une modification plus importante dans ces deux zones. A l'issue de ces observations, les facteurs de modification ont été augmentés davantage afin de faire les différences perceptives entre les sons modifiés encore plus nettes.

12.2.6 Test d'égalisation

Ce test a pour objectif de trouver le facteur d'égalisation de niveau perçu entre les deux types de sons, son 1 et son 2, et des valeurs quantitatives de modification produisant une différence perceptive parmi les 8 sons avec des modifications de zone. Le test est effectué sur 8 personnes : 5 pour le son 1 et 3 pour le son 2. Les sons testés sont 9, dont 8 avec les mêmes modifications (**Fig. 12.11**) que l'exemple précédent.

Consignes Deux tâches sont imposées aux auditeurs. La première consiste à comparer par paire chaque son avec le son sans modification de zone et d'égaliser leur niveau à l'aide d'un curseur dont le pas de précision est de 1 dB. Le son de référence est fixé à 54 dB dans les deux cas. La deuxième tâche consiste à égaliser le niveau des deux sons non modifiés, son 1 et son 2, de deux manières : d'abord en ajustant le son 2, en ayant le son 1 comme référence, ensuite en ajustant le son 1, en ayant le son 2 comme référence.

Résultats Les corrections de niveau de chaque auditeur sont tracées avec des cercles (**Fig. 12.13**). L'abscisse, de 1 à 9, présente les sons modifiés, en commençant avec le son sans modification.

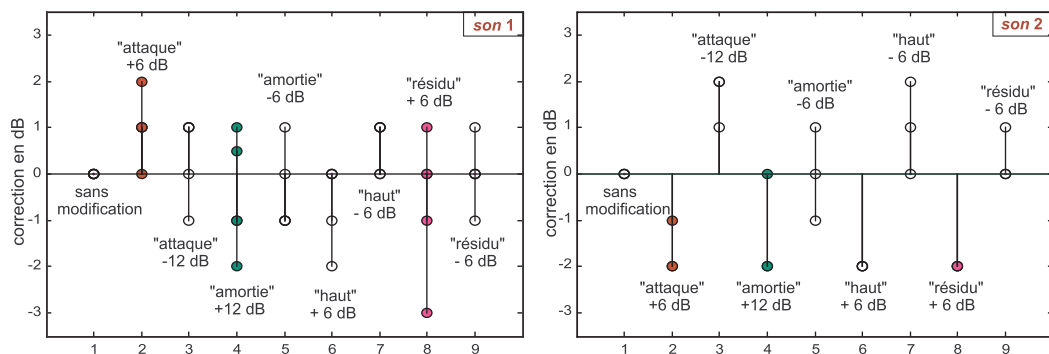


FIG. 12.13 – L'égalisation de sons modifiés par rapport au son sans modification de zone

L'égalisation en niveau du son 1 semble introduire un peu de confusion, car les corrections ne sont pas faites dans la même direction par tous les sujets, comme le montre le son 1 avec modification de +12 dB dans la zone amortie, et + 6 dB de la zone résidu. Les zones amortie et résidu sont donc susceptibles d'influencer la perception de niveau. Pour le son 2, les facteurs sont repartis logiquement, au sens où pour un son avec amplification de zone, les auditeurs diminuent le niveau afin de l'égaliser par rapport au son sans modification, surtout pour l'amplification des zones "haut" et "résidu". L'égalisation des deux sons sans modification a révélé que le son 1 est perçu entre 3 et 4 dB plus fort que le son 2 et ce, quelque soit le son de référence. Cette différence est estimée à environ 6 dB par la podération dB(A), où le niveau du son 1 et son 2 sont déterminés respectivement à 50dB (A) et à 44 db(A) (**Fig. 12.12**).

Conclusion Selon les commentaires des sujets, la tâche d'égalisation est difficile et la précision de 1 dB n'est pas suffisante pour égaliser les bruits. Une précision plus fine que 1 dB est donc nécessaire pour estimer le niveau perçu des sons impulsifs. La modification de 6 dB est insuffisante pour discriminer les sons avec modification dans les zones "haut" et "résidu". Les facteurs d'amplification doivent donc être augmentés. La difficulté d'égaliser les bruits impulsifs en sonie a fait que l'égalisation en niveau n'a pas été entreprise pour les tests de préférence et de catégorisation, présentés plus loin. Il semble que le niveau sonore soit étroitement lié à d'autres facteurs perceptifs comme la hauteur, le timbre et l'impulsivité.

12.3 Outils statistiques

La conception et le traitement des données issues des tests perceptifs s'appuient principalement sur l'utilisation de trois outils statistiques : le plan d'expérience, l'Analyse Multi-Dimensionnelle (MDS) et l'Analyse en Composantes Principale (ACP). Le plan d'expérience se propose d'optimiser l'organisation des essais afin d'obtenir un maximum de renseignements en un minimum d'expériences tout en rendant la meilleure précision sur la modélisation des résultats. La MDS est un algorithme qui permet de reconstruire des distances entre les "produits" à partir des dissimilarités observées et ainsi de connaître les proximités perceptives entre les produits. Enfin, l'ACP est un outil pour réduire les principales dimensions structurant l'espace des produits, tout en assurant l'indépendance des axes principaux.

12.3.1 Plan d'expérience

Nous allons nous référer à la méthode des plans d'expérience afin d'optimiser le nombre d'expériences pour les tests psychoacoustiques. En pratique, cette optimisation consiste à déterminer le nombre total de sons de synthèse et à choisir les modifications à apporter aux signaux.

Processus d'expérience Le procédé d'expérience a pour objectif d'observer un phénomène physique afin de déduire le modèle d'interaction "facteurs-réponse" (Fig. 12.14). Une expérience se propose d'expliquer la réponse expérimentale en fonction des facteurs influant le système. On distingue deux types de facteurs : contrôlés et non contrôlés. Dans un contexte sonore, des facteurs comme le niveau sonore et le timbre influencent la réponse auditive, par exemple, la perception de l'agrément des sons.

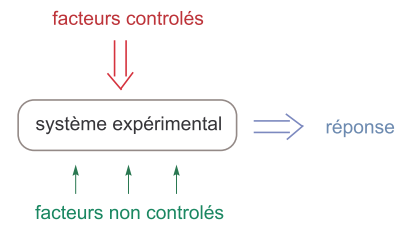


Fig. 12.14 – Le processus d'expérience

Démarche La méthodologie des plans d'expérience aide à la résolution de trois problèmes principaux : la sélection des facteurs influants sur la réponse, la description des variations de la réponse en fonction des fluctuations des facteurs et enfin, la détermination des combinaisons de facteurs pour optimiser la réponse [192]. Deux démarches sont utilisées : déterministe et empirique. L'approche "déterministe" utilise les lois théoriques pour modéliser les phénomènes. L'approche empirique n'a pas d'"a priori" sur le processus et considère le système expérimental comme une boîte noire. Cette deuxième approche, qu'on adopte, va essayer de comprendre le processus à l'aide de plusieurs séries d'expériences.

Formalisme et modèles La démarche des plans expérimentaux pour expliquer un processus dépendant de p facteurs x_1, \dots, x_p qui influencent la réponse expérimentale y , peut être formalisée à l'aide d'un modèle du type :

$$y = f_\beta(x_1, x_2, \dots, x_p) + \epsilon \quad \Longleftrightarrow \quad y = X\beta + \epsilon \quad (12.1)$$

L'objectif de cette modélisation est d'estimer les paramètres du modèle β , tout en minimisant l'erreur ϵ . La fonction f_β est en général inconnue. La méthode traditionnelle statistique de régression linéaire permet l'approximation de f_β à l'aide d'un polynôme. Pour deux facteurs, le modèle de premier degré avec interactions est le modèle le plus simple. Lors qu'il n'est pas suffisant pour décrire le processus, on peut utiliser un modèle de second degré,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon \quad \text{modèle de 1e degré} \quad (12.2)$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \epsilon \quad \text{modèle de 2e degré} \quad (12.3)$$

incluant les effets quadratiques entre les facteurs. Une estimation couramment utilisée de β est donnée par la formule des moindres carrés

$$\beta = (X'X)^{-1} X' y \quad (12.4)$$

Matrice d'information Cette démarche détermine la méthodologie de la surface de réponse y . La surface de réponse est censée représenter les lois de variation de la réponse en fonction des facteurs. La qualité de l'estimation est vérifiée à l'aide de la matrice de variance de β

$$\text{Var}(\beta) = \sigma^2 (X'X)^{-1} \quad (12.5)$$

Par conséquent, la qualité d'estimation des paramètres du modèle va dépendre de deux quantités - l'erreur expérimentale, mesuré à travers σ^2 et la matrice $(X'X)$, appelée souvent matrice d'information. La matrice d'information joue un rôle important dans les critères d'optimalité d'un plan.

12.3.2 Facteurs et modalités

Nous allons nous référer aux plans expérimentaux afin d'optimiser le nombre de facteurs et leurs modalités. La qualité du modèle de prédiction dépend grandement de la construction du plan expérimental. La construction d'un plan passe par une première étape, appelée "screening" ("criblage") des facteurs. Elle comprend la sélection des facteurs influant sur la réponse et le choix des valeurs, ou encore les modalités, que les facteurs vont prendre.

Facteurs et modalités pour les bruits au ralenti Le plan expérimental des tests sur les bruits au ralenti adopte comme facteurs les quatre zones temps-fréquence. Ce sont : Zone 1 ("attaque"), Zone 2 ("amortie"), Zone 3 ("haut") et Zone 4 ("résidu"). Ces 4 facteurs prennent 3 modalités chacun : amplification de zone, atténuation de zone et zone sans modification (**Fig. 12.15**). Les modalités comprennent les valeurs (en dB) d'amplification ou d'atténuation relatives à l'amplitude des atomes dans ces zones.

Les modalités ne prennent pas les mêmes valeurs pour toutes les zones. Ces valeurs sont choisies dans le but d'étudier l'équilibrage de niveau entre les zones et leur influence sur la perception. Le test d'égalisation a montré que des zones à faible énergie comme "résidu" doivent avoir un facteur d'amplification plus grand pour que les sons modifiés soient discriminables. En revanche, la zone "attaque", possédant déjà une très grande énergie, ne nécessite pas une amplification importante. La modification importante de la zone "amortie" de ± 15 dB est faite dans l'objectif d'égaliser son niveau avec la zone "attaque" et d'étudier l'effet perceptif de la succession temporelle entre ces deux zones. La zone "haut" possède un facteur de ± 12 dB afin de tenir compte de la pondération de niveau, plus importante dans les hautes fréquences.

facteur modalité	Zone 1 "attaque"	Zone 2 "amortie"	Zone 3 "haut"	Zone 4 "résidu"
amplification	+ 6 dB	+ 15 dB	+ 12 dB	+ 15 dB
sans modification	0 dB	0 dB	0 dB	0 dB
atténuation	- 15 dB	- 15 dB	- 12 dB	- 15 dB

FIG. 12.15 – Facteurs et modalités choisis pour le plan expérimental des bruits de synthèse

Les valeurs des modalités sont choisies identiques pour les deux sons de synthèse, son 1 et son 2, même si leur répartition énergétique est différente. Le son 1 est plus impulsif, avec une zone "attaque" prédominante et une grande différence entre les zones "attaque"- "résidu" et "attaque"-amortie". Le son 2 possède des niveaux plus équilibrés entre les zones. L'objectif ici est d'observer l'effet perceptif d'une même modification sur deux sons possédant un niveau relatif différent des zones temps-fréquence.

Effets principaux et interactions Dans la théorie des plans expérimentaux, l'influence sur la réponse d'un seul facteur représente un "effet principal". L'effet principal d'un facteur est calculé comme la différence entre les moyennes des réponses pour la modalité la plus et la moins élevée. L'influence sur la réponse d'une combinaison de facteurs est connue comme un "effet des interactions". Dans notre cas, l'interaction est représentée par la modification simultanée de deux ou plusieurs zones temps-fréquence.

12.3.3 Optimalité d'un plan

Un plan expérimental qui contient toutes les combinaisons entre les facteurs et leurs modalités s'appelle un plan complet. C'est un plan qui mène souvent à un nombre d'expériences très élevé. Toutefois, il est possible de construire des plans incomplets, i.e qui utilisent un nombre restreint d'expériences, tout en conservant un maximum d'information sur le procédé et en assurant la fiabilité de la réponse. Notre plan complet de quatre facteurs à trois modalités, demande $3^4 = 81$ expériences. On va rechercher un tel plan optimal capable de réduire le nombre d'expériences et d'estimer en priorité : les effets principaux de chacune des 4 zones temps-fréquence et l'interaction entre la Zone 1 ("attaque") et la Zone 2 ("amortie").

Critères d'optimalité Les critères d'optimalité d'un plan sont basés sur l'étude des propriétés de la matrice d'information ($X'X$). Cette matrice conditionne la qualité des estimateurs et dépend de la matrice d'expérience et le type de modèle choisi. Parmi les critères de qualité d'un plan expérimental sont l'orthogonalité et l'isovariance par rotation. L'orthogonalité exige une matrice de dispersion $(X'X)^{-1}$ diagonale

qui assure l'indépendance des coefficients de régression. La non indépendance des termes quadratiques empêchent l'orthogonalité pour un modèle de deuxième degré, où seulement une quasi-orthogonalité peut être atteinte. Le critère le plus courant est le critère de D-optimalité. Pour une matrice d'effet X , il maximise le déterminant de la matrice d'information $X.X^{-1}$. D'autres critères minimisent la trace de $X.X^{-1}$ (A-optimalité) ou la plus grande valeur de la diagonale (G-optimalité). La qualité d'un plan traite aussi la disposition spatiale des points expérimentaux. Le critère d'isovariance par rotation assure que la distribution de variance du prédicteur de la réponse dépende que de la distance à l'origine de l'espace.

Plan optimal pour les bruits au ralenti En pratique, la construction d'un plan peut se heurter à des difficultés, provenant des contraintes expérimentales et du mélange de facteurs qualitatifs et quantitatifs. Dans ce cas, il existe des algorithmes itératifs de recherche du plan optimal. Ces algorithmes demandent d'imposer au départ l'ensemble des expériences qui correspondent aux contraintes du problème et qui doivent figurer dans le plan final. Dans le cas où cet ensemble n'est pas imposé, l'algorithme peut proposer arbitrairement un ensemble d'expériences.

En se basant sur cette approche, le plan d'expérience final pour les bruits au ralenti a été généré par le logiciel Nemrodw [193]. Le meilleur plan a été calculé, afin d'obtenir une meilleure efficacité et un nombre restreint d'expérience (Fig. 12.16). Il s'est avéré que l'algorithme de calcul du plan optimal évite souvent les combinaisons à un seul facteur, en proposant des plans composés avec des combinaisons de facteurs. Cette approche est justifiée dans la plupart des cas expérimentaux, mais n'est pas adaptée dans notre cas.

Propriété	Valeur
Déterminant ($X'X$)	7.1107
Déterminant (M)	$1.450304E - 0006$
Déterminant	0.355534
Fonct. de variance max	1.060870
Trace ($X'X$) - 1	2.7649
Efficacité G (%)	61.27

Fig. 12.16 – Les propriétés du plan d'expérience pour les bruits au ralenti

Afin de s'adapter à notre expérience, il a été nécessaire d'imposer à l'algorithme de recherche du plan optimal les 9 premières combinaisons. Ces combinaisons comportent les modifications à un seul facteur (8 au total) et le son sans modification. Par la suite, l'algorithme de recherche du plan optimal complète l'ensemble avec d'autres combinaisons de facteurs et modalités afin de garantir son optimalité. Le plan final comporte au total 20 combinaisons (Fig. 12.22). Ce nombre de sons (combinaisons) est raisonnable pour un test de catégorisation, afin de ne pas trop alourdir les séances d'écoute.

12.3.4 Analyse multidimensionnelle (MDS)

L'analyse multidimensionnelle (MDS), introduite par Shepard, est une technique d'analyse des dissimilarités entre des objets [194]. Elle est largement utilisée pour le traitement des données sensorielles. Les dissimilarités entre les objets sont représentées comme des distances entre des points dans l'espace, où chaque point correspond à un objet. La MDS reconstruit une carte des objets à partir de leur matrice de dissimilarité. Cette matrice carrée contient les coefficients de dissimilarité $\Delta = (\delta_{i,j})$, où $\delta_{i,j}$ désigne la dissemblance entre les objets i et j . Les coefficients de dissimilarité prennent une valeur 1 si les objets ont été jugés différents et 0 s'ils sont identiques. La matrice de dissimilarité est symétrique et contient des zéros sur son diagonale (Fig 12.17).

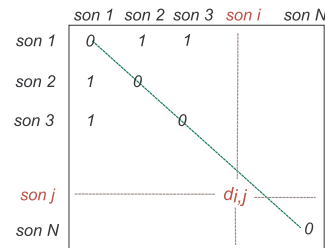


Fig. 12.17 – Matrice de dissimilarité de N sons

La MDS détermine, à partir de la matrice de dissemblance, les coordonnées de n objets dans un espace de p dimensions. Ceci est fait à l'aide de la matrice $X = (x_{il})$ de taille $n \times p$, afin que la matrice des distances entre les points $D(X) = (d_{ij}(X))$ soit la plus proche possible de la matrice des dissemblances Δ , souvent comme la distance euclidienne : $d_{ij}(X) = \sqrt{\sum_{l=1}^p (x_{il} - x_{jl})^2}$.

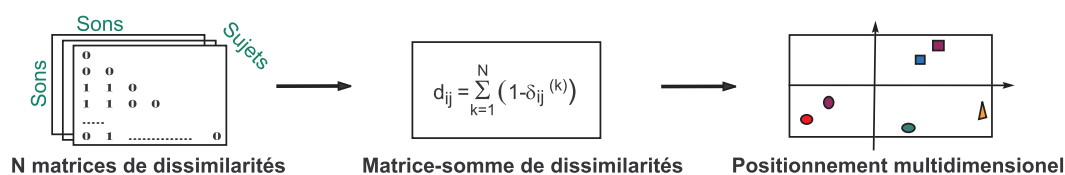


FIG. 12.18 – Démarche de l'analyse multidimensionnelle

La MDS utilise la somme des matrices des dissimilarités qui contiennent les réponses de plusieurs sujets et les dispose sur une représentation plus commune à 2 dimensions (**Fig. 12.18**).

Critère de Stress L'optimisation de la matrice X est basée sur la minimisation itérative du critère de Stress, mesurant la proximité entre $D(X)$ et Δ [195]. Ce critère permet la reconstruction des distances d_{ij} entre les objets les plus proches des dissimilarités (δ_{ij}) observées :

$$Stress(X) = \sum_{i < j} (d_{ij}(X) - \delta_{ij})^2 \quad (12.6)$$

Le critère de Stress détermine également le nombre de dimensions de l'espace multidimensionnelle. Un critère de Stress proche de 0 donne le nombre de dimensions pour donner une meilleure représentation. Le critère de Stress doit être strictement inférieur à 0.1, afin d'avoir une bonne approximation des données en fonction des dimensions choisies.

Nombre de dimensions Le critère de Stress pour les deux familles de son 1 et son 2 qui font l'objet de notre étude, détermine 4 dimensions principales (**Fig. 12.19**). Pour le son 1, on peut choisir également 3 dimensions principales, car cette estimation se trouve au-dessous de 0.1 (10%) de la courbe de Stress. Cependant, il est mieux de prendre une représentation à 4 dimensions, car elle se trouve au-dessous de 0.05 de la courbe de Stress, car un facteur plus proche de 0 donne une meilleure approximation des données expérimentales. L'ajustement des données pour une valeur de Stress choisie est présenté par le diagramme de Shepard. Il montre l'écart du modèle de la réponse expérimentale. Pour les deux types de sons avec des valeurs de Stress respectives de 5.14 % et 4.01 %, on obtient une très bonne corrélation entre les données et le modèle ($R_s > 0.95$).

La MDS reconstruit, à partir des dissimilarités, la carte des distances entre les objets. Cette carte correspond à un positionnement des données sur 2 dimensions choisies. La représentation fidèle des distances est assurée par l'analyse en composantes principales, calculée à partir du résultat de la MDS.

12.3.5 Analyse en composantes principales (ACP)

L'analyse de composante principale est une méthode qui permet l'analyse de données multivariées présentées sous la forme d'un tableau dit "individus-variables" $X = (x_{ij})$ de taille $n \times m$.

Chaque individu (son) représente une ligne du tableau et est décrit par m variables (sujets). L'objectif de l'ACP est d'étudier les similarités entre les individus afin de trouver des individus proches dans l'espace de représentation. D'autre part, l'ACP permet d'étudier les liens éventuels entre les variables, s'il existe des variables corrélées ou des variables qui s'opposent. Le tableau initial de données est vu par l'ACP comme un nuage de points à p -dimensions. L'algorithme recherche les directions de l'espace dans lesquelles la dispersion des données est maximale. Ces directions sont appelées les axes principaux d'inertie.

On utilisera les outils statistiques décrits ci-dessus dans l'analyse des résultats de deux tests : test de catégorisation et test de préférence. Ces tests utilisent 2 familles de 20 sons de synthèse, obtenus à partir des sons de synthèse son 1 et son 2 et à l'aide des modifications de zones selon le plan d'expérience (**Fig. 12.22**). Les sons de type 1 et de type 2 sont respectivement sur (**Piste 143-162** et **Piste 163-182**).

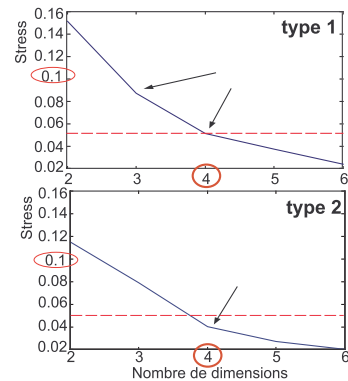


Fig. 12.19 – Critère de Stress : sons type 1 et sons type 2

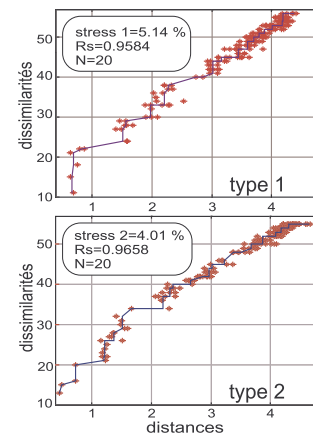


Fig. 12.20 – Le diagramme de Shepard pour les sons de type 1 et type 2

	sujet 1	sujet 2	sujet 3	sujet j	sujet M
son 1	x_{11}	x_{12}	x_{13}		
son 2	x_{21}				
son 3	x_{31}				
son i				x_{ij}	
son N					x_{NM}

Fig. 12.21 – Matrice des "individus-variables" pour l'ACP


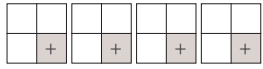


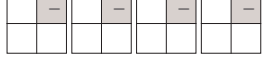



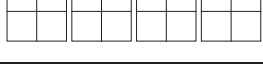

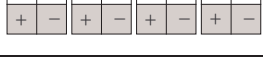









N son	Description	Schéma	Zone 1	Zone 2	Zone 3	Zone 4
Son 1	amortie -15 dB		0	- 15 dB	0	0
Son 2	amortie +15 dB		0	+15	0	0
Son 3	haut -12 dB		0	0	-12	0
Son 4	haut +12 dB		0	0	+12	0
Son 5	résidu -15 dB		0	0	0	-15
Son 6	résidu +15 dB		0	0	0	+15
Son 7	attaque -15 dB		-15	0	0	0
Son 8	attaque +6 dB		+6	0	0	0
Son 9	normal		0	0	0	0
Son 10	attaque -15 dB ; amortie + 15 dB		-15	+15	0	0
Son 11	attaque +6 dB ; amortie -15 dB		+6	-15	0	0
Son 12	attaque - 15 dB ; haut +12 dB		-15	0	+12	0
Son 13	attaque -15 dB ; résidu + 15 dB		-15	0	0	+15
Son 14	attaque - 15 dB ; amortie - 15 dB		-15	-15	0	0
Son 15	attaque + 6 dB ; amortie + 15 dB		+6	+15	0	0
Son 16	amortie - 15 dB ; haut - 12 dB		0	-15	-12	0
Son 17	haut - 12 dB ; amortie - 15 dB		0	0	-12	-15
Son 18	attaque +6 dB ; haut + 12 dB		+6	0	+12	0
Son 19	amortie - 15 dB ; résidu - 15 dB		0	-15	0	-15
Son 20	amortie + 15 dB ; résidu + 15 dB		0	+15	0	+15

FIG. 12.22 – Le plan d'expérience retenu pour les modifications à 4 zones et 3 modalités

12.4 Test de catégorisation

Les deux sections suivantes présentent les tests effectués avec l'ensemble des bruits de synthèse de ralenti diesel : test de catégorisation et test de préférence. Le début de cette section sera consacré à la procédure de déroulement des tests perceptifs. Par la suite seront exposés les données expérimentales du test de catégorisation et du test de préférence. Une discussion sur les résultats et les perspectives conclura la présente thèse.

12.4.1 Protocole des tests

Cette section décrit la préparation générale des tests : la construction des sons de synthèse, l'organisation des séances, les auditeurs, et les consignes pour les test de préférence et de catégorisation. A la fin sont résumés les objectifs des tests perceptifs.

Sons de synthèse Des sons de synthèse sont obtenus à partir de deux sons originaux, exprimant des réglages différents d'un moteur diesel : sans injection pilote (**Piste 104**) et avec injection pilote (**Piste 106**). L'analyse-synthèse d'un cycle moteur des ces deux sons a été présenté à la fin de la deuxième partie de la thèse (**Fig.10.32**). Les sons de synthèse respectifs, le son 1 (**Piste 105**) et le son 2 (**Piste 104**) sont à la base de deux ensembles de 20 sons de synthèse, qu'on appelle des sons de type 1 et des sons de type 2. Les sons de chaque ensemble sont obtenus en amplifiant ou atténuant l'amplitude des atomes dans une ou deux zones temps-fréquence, d'après le découpage illustré sur la (**Fig. 12.4**) et selon le plan d'expérience résumé sur la (**Fig. 12.22**). Les ensembles de sons de type 1 et de type 2 sont présentés séparément dans deux séances, car ces deux types de sons possèdent des timbres et des niveaux différents, ce qui les rend difficilement comparables.

Niveau sonore Le niveau sonore des sons de synthèse sans modification a été fixé à 54 dB. Ce niveau est certainement inférieur au niveau réel des bruits enregistrés en banc moteur, mais représente un niveau optimal pour laisser les auditeurs évaluer les bruits sans éprouver une gêne auditive trop importante. Le niveau de 54 dB les différences de niveau entre son 1 et son 2, où le son 1 est perçu plus fort que le son 2, comme le montrent la pondération dB(A) (**Fig. 12.12**) et le test d'égalisation, présentés précédemment. Les sons de synthèse présentés ne sont pas égalisés préalablement, d'un côté à cause de la tâche relativement difficile d'égalisation et d'un autre côté, pour observer directement l'influence du niveau sonore sur les évaluations perceptives des auditeurs.

Déroulement des tests Les séances d'écoutes se sont déroulées séparément pour l'ensemble des sons de type 1 et de type 2. Ces séances ont commencé par le test de préférence afin de familiariser les auditeurs avec la nature des sons à évaluer et ont poursuivi avec le test de catégorisation. Les deux séances de tests sont organisées comme suit

Séance "sons de type 1" : Test de préférence (type 1) + test de catégorisation (type 1).
Séance "sons de type 2" : Test de préférence (type 2) + test de catégorisation (type 2).

Le test s'est déroulé entre 35 et 50 min, dont en moyenne 10 minutes pour le test de préférence et 30 minutes pour le test de catégorisation. Le test de préférence a lieu en premier afin avoir une réponse spontanée des auditeurs sur leurs sensations hédoniques et de préparer. En effet, la possibilité de réécoute donnée lors de la catégorisation peut rendre moins spontanées les réponses hédoniques des auditeurs si le test de préférence se déroule après le test de catégorisation. L'ordre de présentation des sons pour le test de préférence est obtenu à l'aide d'un carré latin qui assure la nature aléatoire de présentation.

Auditeurs Dans les séances de tests ont participé 111 personnes, dont 28 femmes et 83 hommes, d'âge compris entre 20 et 60 ans. Ces auditeurs n'ont pas déclaré de déficience auditive. Parmi eux, 56 ont passé la séance d'évaluation de sons de type 1, 55 ont évalué des sons de type 2, où un minimum de 55 personnes est exigé pour la validation statistique des résultats. La méthodologie de tests perceptifs exige que pour la séance de chaque type de sons, type 1 ou type 2, au moins la moitié des sujets participent pour la première fois à ce type de test.

Consignes pour le test de préférence Les auditeurs sont d'abord informés que les sons qu'ils vont écouter représentent des bruits de synthèse de ralenti diesel. Afin de familiariser les sujets avec la nature des sons à évaluer, quatre sons sont mis à l'écoute au début de la séance. Les sujets sont ensuite invités à donner leur appréciation pour chaque son, sur une échelle allant de "je n'aime pas" à "j'aime bien" (**Fig.**

12.23, à gauche). Ces termes qualitatifs sont estimés plus neutres que l'adjectif "agréable" qui demande un engagement hédonique difficilement appropriable aux sons impulsifs répétitifs. L'expression en négation "plus désagréable-moins désagréable" n'est pas envisageable car, selon la méthodologie des tests, la double négation a tendance de perturber les repères d'évaluation des auditeurs. L'échelle d'évaluation du test de préférence code les réponses des sujets sur une plage continue de 0 à 10. Les auditeurs ont la possibilité de réécouter les sons, mais ils ne peuvent plus changer d'avis une fois leur évaluation donnée.

Consignes pour le test de catégorisation Le test de catégorisation suit immédiatement le test de préférence pour les sons du même type. Les auditeurs se retrouvent devant une interface exposant 20 sons présentés avec des numéros codés permettant de faire la correspondance avec les vrais sons (**Fig. 12.23, à droite**). La consigne aux sujets a été d'écouter attentivement les 20 sons présentés et de les regrouper selon leurs ressemblances ou de séparer les sons différents dans des groupes différents.

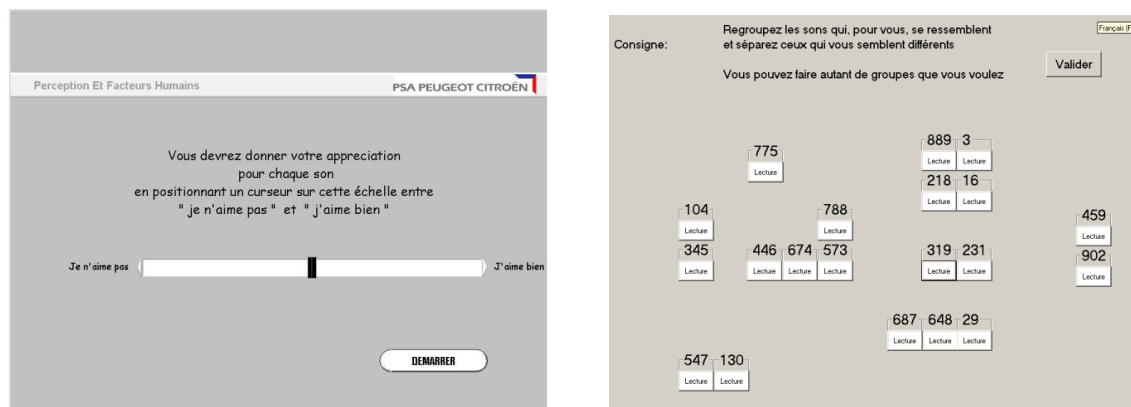


FIG. 12.23 – Interfaces pour le test de préférence (à gauche) et le test de catégorisation (à droite)

Le critère de regroupement, le nombre des groupes et le nombre de sons par groupe sont laissés au choix de l'auditeur qui peut, s'il le souhaite, réécouter les sons autant fois que nécessaire. Après la validation des groupes formés, les participants sont invités à noter sur une feuille les sons de chaque groupe et à décrire librement avec des mots leurs particularités. A la fin, ils sont demandé de classer globalement les groupes en ordre de préférence.

Objectifs Les tests perceptifs ont pour objectif de démontrer l'influence des zones temps-fréquence sur la perception de deux types de bruits de synthèse, correspospondant aux réglages différents du moteur diesel. Le test de catégorisation permettra de faire ressortir sur les principales dimensions les indicateurs perceptifs utilisés par les auditeurs pour distinguer les bruits de synthèse. Plus particulièrement, notre objectif sera de démontrer l'existence des flux auditifs, une notion de rythme et un "équilibre" entre les zones temps-fréquence pris en compte dans l'évaluation perceptive des bruits impulsif répétitifs. Le test de préférence donnera l'appréciation des auditeurs pour les sons à l'intérieur de chaque type.

12.4.2 Présentation des dimensions

Les résultats du test de catégorisation sont résumés dans une matrice de dissemblance qui est analysée par le procédé de la MDS (**Fig. 12.18**) afin de donner la représentation des données sur 4 dimensions principales, déterminées par le diagramme de Stress (**Fig. 12.19**).

Inertie des dimensions L'inertie est une mesure de l'importance de chaque dimension dans l'explication des données de la catégorisation est déterminée par le pourcentage d'inertie qui lui est associée. Le pourcentage d'inertie pour les deux types de sons sur les 4 dimensions est répartie comme suit :

Type 1	Pourcentage d'inertie	Type 2	Pourcentage d'inertie
Dim 1	62.65 %	Dim1	74.15 %
Dim 2	18.75 %	Dim 2	17.75 %
Dim 3	10.48 %	Dim 3	5.31 %
Dim 4	8.12 %	Dim 4	2.79 %

Au total, les quatre dimensions décrivent à 100 % les données pour les deux types de sons. Le pourcentage d'inertie montre que les résultats de la catégorisation sont explicables à 90 % par les 3 premières dimensions pour les sons de type 1, les 2 premières dimensions pour les sons de type 2. On verra plus tard que la 3e dimension est associée à l'évaluation des préférences des sujets pour les sons de type 1.

Représentation des données par la MDS La représentation des résultats montre que les auditeurs ont effectué un regroupement relativement similaire des sons sur les 2 premières dimensions, au sens où les mêmes modifications ont provoqué des différences perceptives similaires pour les deux types de sons (**Fig. 12.24**). En revanche, les dimensions 3 et 4 sont différentes pour les deux types de sons. Afin de mieux situer les sons modifiés, le son 9, sans modification, est présenté en vert.

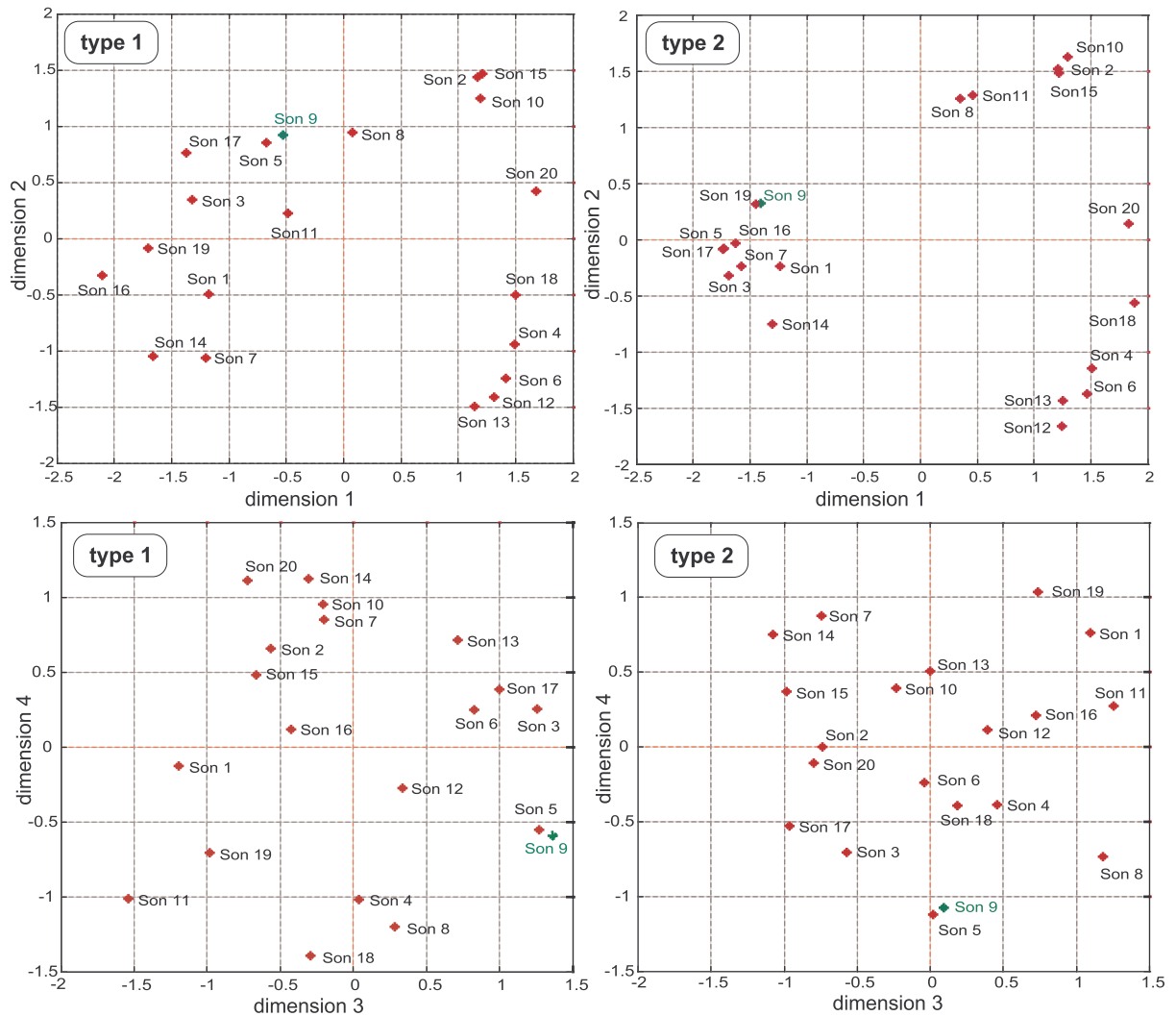


FIG. 12.24 – Résultats de la catégorisation : projection de l'espace MDS sur les dimensions 1 et 2, et sur les dimensions 3 et 4 les sons de type 1 et type 2

Groupes de sons Le résultat le plus représentatif du regroupement effectué par les sujets est donné par le compromis statistique. Le compromis détermine les sons qui ont été le plus souvent regroupés ensemble. Ces groupes sont repartis comme suit

Type 1 : {2,10,15}, {1,19}, {3,17}, {6,12,13}, {5,9}, {7,14}, {4,18}, {8}, {11}, {16}, {20}
Type 2 : {2,10,15}, {1,19}, {3,17}, {12,13}, {5,9}, {7,14}, {4,6}, {8,11}, {16}, {18}, {20}

Ces groupes vont permettre de mieux interpréter les dimensions et le critère de regroupement utilisé par les auditeurs.

12.4.3 Effets des zones sur les dimensions

L'influence de chacune des zones selon le plan d'expérience peut être étudié à l'aide de l'analyse de la variance (ANOVA). Elle donne une information quantitative sur le poids relatif de chacun des facteurs (zones), et permet d'estimer, d'un point de vue statistique, lesquels sont les facteurs influant sur les différentes dimensions. L'effet des zones sur les dimensions est résumé sur le tableau suivant, où les zones influant sur une dimension sont citées par ordre d'importance :

Type 1 :

Dim 1 : Zone 2, Zone 3; Zone 4;

Dim 2 : Zone 2; Zone 4; Zone 3; Zone 1

Dim 3 : Zone 2

Dim 4 : Zone 1, Zone 2, Zone 3; Zone 1*Zone 2

Type 2 :

Dim 1 : Zone 3; Zone 4; Zone 1;

Dim 2 : Zone 2, Zone 3, Zone 4; Zone 1; Zone 1*Zone 2;

Dim 3 : Zone 2

Dim 4 : (Zone 2)

Cette répartition des effets témoigne de la complexité d'interaction entre les zones et leur influence sur chaque dimension. Ces données statistiques peuvent s'avérer également utiles dans l'interprétation des dimensions. En effet, la disposition des zones sur les dimensions 1 et 2 (**Fig. 12.25**) permet de faire une équivalence avec la disposition des sons sur les mêmes dimensions.

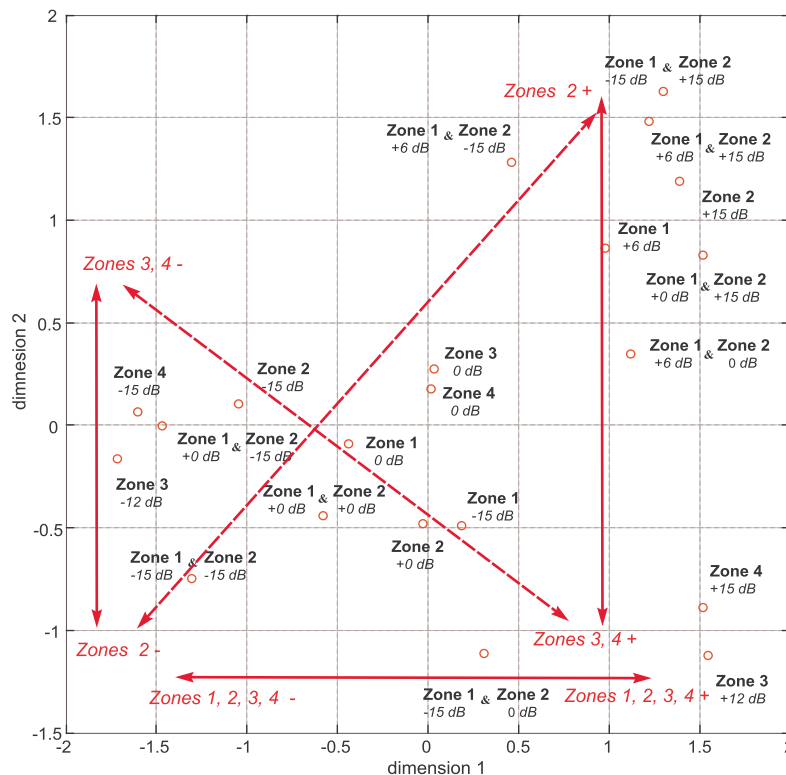


FIG. 12.25 – Disposition des zones temps-fréquence sur les dimensions 1 et 2

Cette disposition des zones permet de donner une interprétation des dimensions 1 et 2. Plus particulièrement, on peut observer que la dimension 1 est associée à une amplification du niveau global des sons : sur l'axe positif (à droite) sont situés des sons possédant une amplification de zones; sur l'axe négatif (à gauche) sont les sons avec une atténuation de zone. La deuxième dimension est liée étroitement avec l'aspect grave-aigu des flux auditifs. L'axe positif (en haut) est associé à la fois avec l'amplification de la zone 2 ("amortie") et l'atténuation des zones 3 et 4 ("haut" et "résidu"), ce qui exprime dans les deux cas la prépondérance d'un flux auditif dans les basses fréquences. Sur l'axe négatif de la dimension 2 on observe une atténuation de la zone 2 et une amplification des zones 3 et 4, ce qui mène à une amplification globale des hautes fréquences. Il est intéressant de noter, qu'il existe un axe en diagonale, qui semble avoir une grande influence dans la catégorisation des sons, et il correspond à l'amplification et l'atténuation de la zone 2 ("amortie"). Plus particulièrement, cet axe semble tenir compte de l'équilibre entre les niveaux des zones "attaque" et "amortie", ce qui exprime autrement l'impulsivité du bruit.

12.4.4 Interprétation des dimensions

Cette section est consacrée à l'analyse des résultats des tests fournis par la MDS. Je présente ici mon interprétation des dimensions perceptives des bruits de synthèse au ralenti, ainsi que mes observations sur plusieurs phénomènes de formation de flux auditifs et de masquage temps-fréquence entre les zones.

Dimension 1 Afin de donner une interprétation de la dimension 1, on peut observer que les sons sont étalés en deux nuages : à droite et à gauche de l'axe central (**Fig. 12.26**). Le nuage à droite contient les sons 2, 4, 6, 8, 10, 12, 13, 15, 18 et 20 et semble disposé d'une manière similaire pour le type 1 et le type 2. En revanche, le nuage à gauche, qui contient les autres sons dont le son sans modification, le son 9, s'étale d'une manière plus large pour les sons de type 1 que pour les sons de type 2. Ceci suggère que les mêmes modifications de zones ont apporté une différence perceptive plus importante pour les sons de type 1 que pour ces sons de type 2.

Les sons appartenant au nuage de droite (2, 4, 6, 8, 10, 12, 13, 15, 18 et 20) possèdent en commun une amplification globale du niveau sonore par rapport au son sans modification. Cette amplification globale est due soit à une amplification d'une seule zone (sons 2, 4, 6 et 8) ; soit à une amplification à la fois de deux zones (sons 15, 18, 20) ; soit à une atténuation de la zone "attaque", combinée avec une amplification d'une autre zone (sons 10, 12, 13) ; soit l'amplification de la zone "attaque", combinée avec atténuation de la zone "amortie". Le résultat dans tous les cas est une amplification globale du niveau sonore. D'une manière similaire, les sons situés dans les nuages à gauche possèdent une atténuation de zones, et par conséquent, une diminution du niveau global perçu.

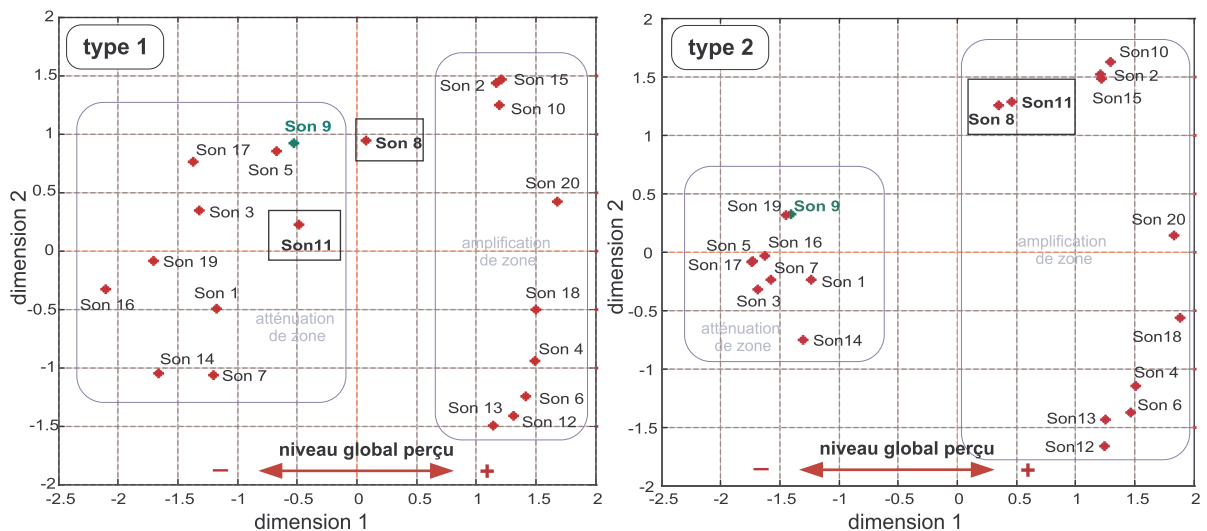


FIG. 12.26 – Interprétation de la dimension 1 comme le niveau global perçu

Deux sons méritent ici un intérêt particulier : le son 8 et le son 11, respectivement avec une amplification de la zone "attaque" à 6 dB et une amplification de la zone "attaque" combinée avec une atténuation de la zone "amortie". On peut apercevoir que pour le type 1, le son 8 est beaucoup plus proche du son sans modification, le son 9, et peut être même attribué au nuage de gauche. Pour le type 2, le son 8 est davantage éloigné du son 9, mais se rapproche du son 11. Cette constatation montre que si la différence de niveau entre les zones "attaque"-"amortie" n'est pas grande (sons de type 2), l'amplification de la zone "attaque" augmente davantage cette différence et rend le bruit plus impulsif (son 8 de type 2), alors que dans le cas où la zone "attaque" possède une grande énergie au départ, son amplification ne change pas considérablement le niveau et l'impulsivité du bruit (son 8 du type 1). Pour le type 1, une atténuation de la zone "amortie" combinée avec une amplification de la zone "attaque" n'a pas de conséquences importantes sur le niveau, car la zone "attaque" est suffisamment énergétique.

Dimension 2 La dimension 2 semble liée à l'aspect "grave-aigu" des sons. Pour les deux types, le regroupement oppose sur l'axe de la dimension 2, le groupe de sons 2, 10 et 15 avec le groupe des sons 4, 6, 12 et 13 (**Fig. 12.27**). Les sons du groupe (2, 10 et 15) ont une amplification de la zone "amortie" qui est une zone relativement forte en basses fréquences. Les sons (4, 6, 12, et 13) ont une amplification dans les hautes fréquences, soit de la zone "haut", soit de la zone "résidu". Les sons 1, 7 et 14 ont une atténuation dans les basses fréquences ("attaque" ou "amortie"), ce qui laisse ressortir davantage les hautes fréquences.

Ces sons reflètent un changement perceptif attendu, dans le sens où une amplification dans les basses fréquences et une atténuation dans les hautes fréquences privilégient la perception des basses fréquences et de même pour les hautes fréquences.

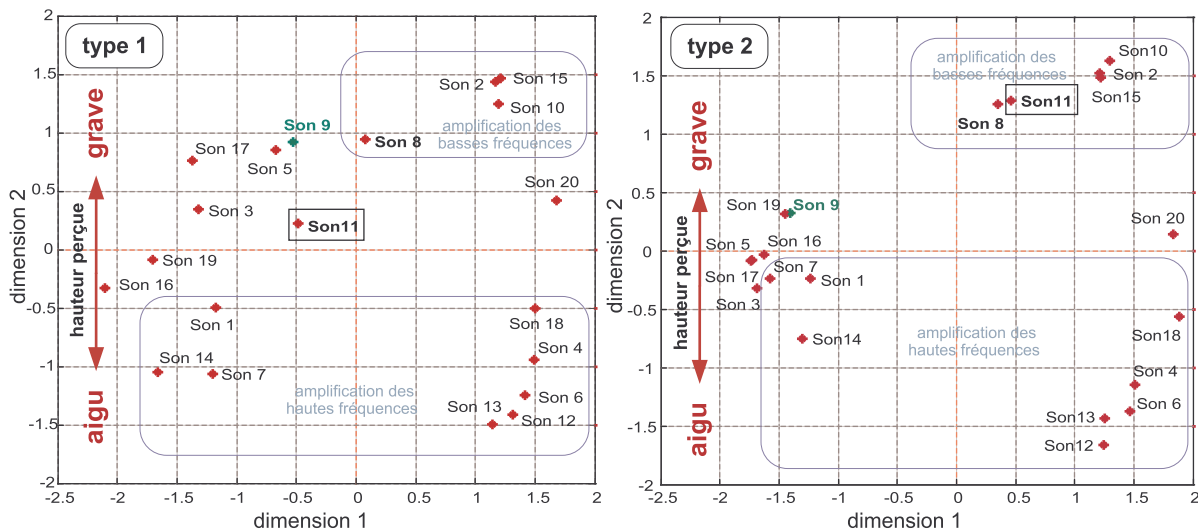


FIG. 12.27 – Interprétation de la dimension 2 comme la hauteur perçue

Les sons qui ne sont pas regroupés sur la (Fig. 12.27) représentent des cas très intéressants, car ils illustrent la situation inverse, où une atténuation dans les hautes fréquences peut accentuer perceptivement davantage les hautes fréquences. Ces phénomènes semblent apparaître suite à la formation des flux auditifs et du masquage temps-fréquence effectué entre les zones. Ainsi, le son 3 possède une atténuation de la zone "haut" et pourtant, il est estimé plus aigu que le son sans modification, le son 9. Probablement, le fait de diminuer la zone "haut" fait ressortir davantage un flux auditif masqué auparavant, formé par la zone "résidu", voisine à la zone "haut" dans les hautes fréquences. Beaucoup moins marquée, la situation inverse où la zone "résidu" est atténuée (son 5) provoque aussi une accentuation vers les aigus, mais la faible énergie de la zone "résidu" fait que les changements perceptifs ne sont pas aussi bien exprimés.

Ces phénomènes de masquage peuvent être détectés chez les sons 16 et 19 qui subissent une atténuation à la fois de deux zones : l'une est la zone "amortie" et l'autre est respectivement la zone "haut" et la zone "résidu". Probablement, ces changements favorisent l'apparition du flux constitué de la zone "résidu" pour le son 16 et en accentuant le flux "haut" pour le son 19. Enfin, l'atténuation à la fois des zones "haut" et "résidu" pour le son 17, représentant une atténuation des hautes fréquences sur toute la plage fréquentielle, est également associée à une accentuation du flux dans les hautes fréquences, beaucoup plus exprimée pour le type 2 que pour le type 1.

Enfin, le son 11, représentant une amplification de la zone "attaque", combinée avec une atténuation de zone "amortie", provoque des changements perceptifs qui dépendent du type de sons, type 1 ou type 2, aussi bien pour le niveau que pour la hauteur perçue. On peut observer que pour le type 1, le son 11 est perçu nettement plus aigu, alors que pour le type 2, il est perçu beaucoup plus grave (Fig. 12.27). Il s'agit de deux effets perceptifs différents qui dépendent fortement de l'équilibrage des niveaux entre les zones pour les deux types. Comme on l'a déjà démontré, le niveau relatif des zones pour les deux types est différent (Fig. 12.10). Le type 1 possède une grande énergie dans la zone "attaque" et une grande différence entre les niveaux des zones "attaque" et "amortie". Le type 2 possède une énergie moins importante dans la zone "attaque" et un niveau équilibré entre les zones "attaque" et "amortie". Ainsi, le fait d'augmenter la zone "attaque" pour le son 11 du type 1 ne change pas beaucoup l'impulsivité du bruit, et l'atténuation simultanée de la zone "amortie" ne fait que ressortir davantage les hautes fréquences, surtout de la zone "haut". Pour le son 11 du type 2, l'amplification de la zone "attaque" est beaucoup trop importante, comme le manifeste par ailleurs le son 8, possédant une amplification de la seule zone "attaque". Ceci provoque une augmentation des basses fréquences pour le son 11 de type 2, où la zone "attaque" se retrouve dans le rôle de zone effectuant un masquage sur les autres zones. Cet effet semble également augmenter l'impulsivité du bruit.

Dimension 3 L'interprétation des dimensions 3 et 4, comme le témoigne la Fig. 12.24, reste moins évidente, à cause de la multitude de facteurs qui entrent en compte dans la structure des bruits : niveau

relatif initial des zones, contenu spectral de la zone, type de modification. Afin de simplifier l'analyse de la dimension 3, on va la représenter en combinaison avec la dimension 1 (**Fig. 12.28**), une dimension connue, puisqu'elle a été déjà interprétée ci-dessus comme liée au niveau global des bruits. La dimension 3 semble être liée au tempo relatif du flux formé dans les hautes fréquences. D'une manière globale, cette dimension révèle certains aspects rythmiques dans la structure fine des bruits.

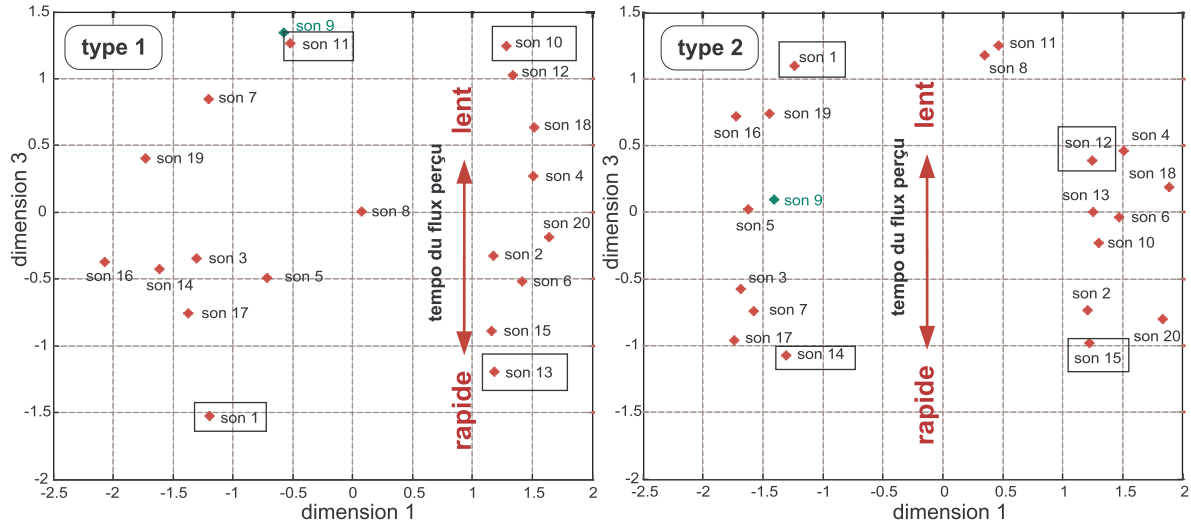


FIG. 12.28 – Représentation des dimensions 1 et 3 pour les type 1 et 2

L'idée du tempo des flux auditif va être illustrée à l'aide des sons positionnés à l'extrême sur l'axe de la dimension 3. Les couples de sons "son 11-son 1" et "son 10-son 13" pour le type 1, ainsi que "son 1-son 14" et "son 12-son 15" pour le type 2 vont servir d'exemple afin d'analyser leur structure rythmique. D'une manière globale, un tempo rapide est créé quand deux zones voisines en temps, comme les zones "haut"- "résidu" ou "attaque"- "amortie", ont un niveau sonore relativement similaire. Par exemple, le son 13 du type 1 possède une énergie importante dans les zones "attaque" et "haut". Le fait amplifier la zone "résidu" remet à un niveau similaire les zones qui se succèdent dans le temps dans les hautes fréquences, "haut" et "résidu". De ce fait, le flux auditif créé dans les hautes fréquences est perçu comme plus rapide (**Fig. 12.29**). En même temps, le son 13 a subi une atténuation de la zone la plus forte, "attaque", ce qui égalise les niveaux des zones "attaque" et "amortie" et crée une sensation d'événements sonores qui se succèdent à un tempo plus rapide.

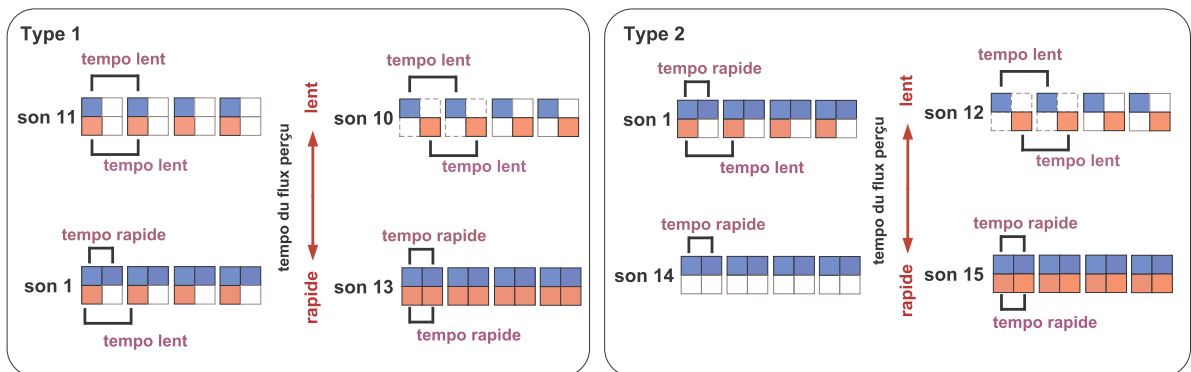


FIG. 12.29 – Formation de flux auditifs, notion de tempo rapide et lent

Il faut noter que les tempo du flux basses fréquences et du flux hautes fréquences vont entrer en compétition et c'est l'auditeur qui décide comment évaluer le tempo global du son. D'une manière générale, il semble que les sujets ont privilégié le tempo du flux perçu plus fort en niveau sonore. Par exemple, dans la présence d'un flux lent dans les basses fréquences et un flux rapide dans les hautes fréquences, pour le son 1 du type 1, c'est le flux haut qui est plus fort, alors que pour le son 1 du type 2, c'est le flux bas qui est privilégié.

Cette troisième dimension est moins évidente à interpréter que les dimensions 1 et 2, associées respectivement au niveau et à la hauteur, et qui semblent plus familières pour l'auditeur. Toutefois, la dimension 3, liée au tempo des séquences sonores dans les basses et dans les hautes fréquences est aussi très importante dans la perception des bruits impulsifs répétitifs. Plusieurs auditeurs se sont exprimés dans la description des sons sur les divers aspects rythmiques comme le tempo, la rapidité, la régularité des séquences. De plus, pour les sons de type 1, la dimension 3 est la deuxième dimension la plus importante dans l'appréciation des bruits, comme le montre le test de préférence, présenté un peu plus loin. Notre objectif de démontrer l'importance de cette dimension liée au rythme et le tempo des bruits au ralenti, dans le contexte des flux auditifs, est désormais atteint.

Dimension 4 Je n'ai pas trouvé une interprétation valable pour cette dernière dimension, qui semble rassembler tous les facteurs "résiduels" qui ne peuvent pas être expliqués par les trois dimensions cités ci-dessus. La dimension 4 contient respectivement 8,12 % pour le type 1 et 2,79 % pour le type 2 de pourcentage d'inertie. Il s'agit donc d'une dimension moins importante, qui n'affecte pas beaucoup la catégorisation, et qui n'influence pas vraiment la préférence des auditeurs.

12.4.5 Verbalisation

Les auditeurs ont eu le libre choix de décrire les groupes de sons qu'ils ont formé lors du test de catégorisation. La difficulté dans l'analyse des résultats de la verbalisation consiste dans la multitude de termes employés.

Niveau sonore La description des sons comporte des termes directs concernant le niveau sonore des bruits comme "fort", "intense", "doux", mais il semble que les auditeurs se sont concentrés beaucoup plus sur le timbre, la hauteur, le rythme et l'impulsivité.

Aspect "grave-aigu" L'aspect "grave-aigu" est décrit par les termes "grave", "aigu" et "médium" et une grande majorité des auditeurs ont fait leur catégorisation en se basant sur les différences de hauteur perçue. Le terme "grave" est très souvent accompagné par "sourd", "feutré" et "bourdonnement", alors que "aigu" est souvent en combinaison avec "claquement", "martèlement", "métallique" et "sifflement". Une partie des sujets ont décéléré la présence de plusieurs flux, ont remarqué cela et ont fait leur regroupement en fonction des différences dues à la fois du flux grave, désigné par "bruit grave" ou "bourdonnement grave" et le flux aigu, nommé "claquement aigu".

Notion de rythme Sur 111 écoutes (certains sujets ont passé les deux tests, 28 évoquent et utilisent dans la description des sons, la notion de rythme. Ceci représente en pourcentage 25 % des sujets. Ils utilisent la notion de "rapide" (en combinaison de "plus", "moins" ou "très") et plus rarement "lent" et "mou". Les descriptions comme "régulier", "irrégulier", "rythme" sont aussi utilisées.

Notion d'impulsivité L'impulsivité des sons est décrite par des mots comme "claquement", très souvent en combinaison avec "aigu" et "métallique" ou "rapide". Autres mots pour décrire sont "impact", "choc", "toc-toc", "cliquetis", "clac", "clapotement", en combinaison avec "saccadé" ou "haché".

Comparaisons Les auditeurs se sont servis de description à l'aide d'appariement aux autres objets. Les plus souvent employés étaient "tracteur" (ou petite machine agricole), "hélicoptère", "bateau", "marteau-piqueur" et "mitrailleuse laser". Il a eu une catégorie relative aux sons "synthétiques", "numérique", "techno" et qui sonnaient d'une manière "artificielle". Episodiquement, ils ont comparé les bruits aux outils comme "moulin à café", "machine à coudre", "machine à écrire", "tambour". Troisième classe d'appariement était les sons de la nature ("chant d'oiseaux", "pie vert", "cigale" et il a eu très particulièrement une comparaison avec de "l'eau qui coule" ("liquide", "bulles", "fluide").

Experts moteur La verbalisation de quelques spécialistes moteur a été différente des auditeurs "naïfs". Les spécialistes ont raisonné en termes de harmonique moteur, en cherchant à noter la présence ou l'absence d'une harmonique.

12.5 Test de préférence

Le test de préférence fait partie des tests perceptifs dits hédoniques. L'importance du test de préférence réside dans le caractère de jugement spontané de la part des auditeurs sur l'agrément ou le désagrément perçu pour chacun des sons présentés. En revanche, le test de préférence ne renseigne pas sur les critères utilisés par les auditeurs. C'est pourquoi le test de catégorisation s'avère complémentaire du test de préférence.

12.5.1 Cartographie des préférences

Les résultats du test de préférence sont illustrés à l'aide de deux représentations : la cartographie interne et la cartographie externe des préférences. La cartographie interne utilise une analyse en composante principale, alors que la cartographie externe prend en compte les dimensions extraites par la MDS [196].

Cartographie interne La matrice de préférence est formée par les évaluations de chaque sujets (en ligne) pour chaque son (en colonne) (**Fig. 12.30**). Cette matrice est reprise pour l'analyse en composante principale (ACP). L'ACP effectuée est non-normée, afin de respecter l'étendue d'échelle, utilisée par chaque auditeur. A l'issue de l'analyse, on obtient une représentation dans un plan, déterminé par les deux premières composantes de l'ACP. Cette graphique est connue comme la cartographie interne de préférence (**Fig. 12.31**)

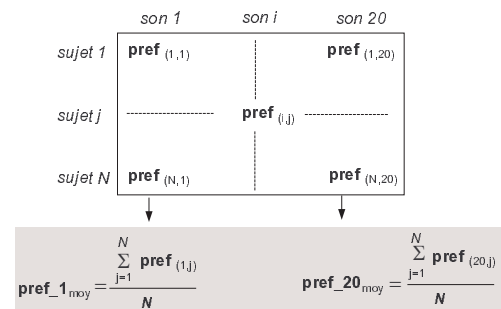


Fig. 12.30 – Matrice de préférence

La cartographie interne permet d'indiquer les sons préférés et les sons rejetés par les auditeurs. Cette observation se fait en fonction de l'axe d'inertie le plus important. La limite entre des sons préférés et rejetés donné sur la (**Fig. 12.31**) est donnée à titre indicatif. Il semble que l'appréciation des auditeurs est presque exclusivement basée sur le critère du niveau global des bruits. En effet, les sons rejetés comportent une amplification de zone, alors que les sons préférés ont une atténuation de zone.

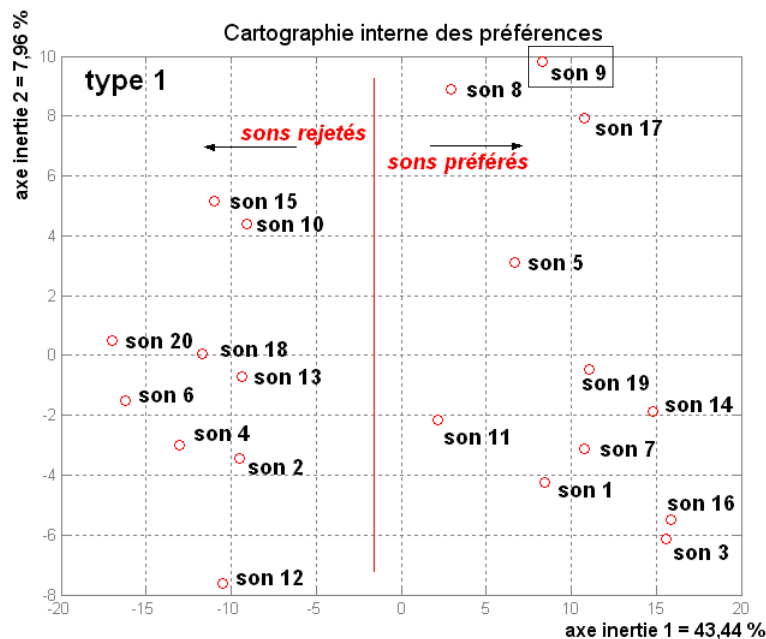


FIG. 12.31 – Cartographie interne de préférence pour les sons de type 1

La relation entre le niveau sonore et la préférence est confirmée plus loin par la cartographie externe

des préférences qui met en relation les résultats de test de préférence avec ceux du test de catégorisation.

Cartographie externe La cartographie externe de préférences intègre les résultats de la catégorisation. Sur ce point, elle diffère de la cartographie interne qui, quant à elle, montre les résultats seulement dans le contexte des préférences. La cartographie externe de préférences se propose de trouver un modèle qui approche bien les données expérimentales. Ce modèle est sous la forme d'une combinaison linéaire des dimensions $D_1...D_n$, trouvées par la catégorisation. Le modèle le plus simple, le modèle linéaire vectoriel, approche les réponses des sujets présentées sous la forme vectorielle Y , avec une combinaison linéaire de deux premières dimensions D_{n_1} et D_{n_2} , déterminées par la catégorisation $D_1...D_n$. Le modèle linéaire est

$$Y = a_0 + a_1.D_{n_1} + a_2.D_{n_2} \quad (12.7)$$

où a_0 est une constante et a_1 et a_2 sont les coefficients du modèle. Le vecteur Y comporte les notes moyennes de tous les sons, $Y = (pref1_{moy}, ..., pref20_{moy})$. Géométriquement, la cartographie externe correspond à un hyperplan. Le modèle vectoriel est un cas particulier du modèle complet,

$$Y = a_0 + \sum_i a_i.D_{n_i} + \sum_i b_i.D_{n_i}^2 + \sum_{ij} c_{ij}.D_{n_i}.D_{n_j} \quad (12.8)$$

D'autres cas particuliers du modèle complet sont le modèle circulaire (les b_i sont égaux et les c_{ij} sont nuls) et le modèle elliptique (les c_{ij} sont nuls). Les résultats de ces modèles se représentent sur une hypersurface quadratique. On adopte le modèle vectoriel (12.7) pour sa simplicité. Il propose l'interprétation la plus intuitive sous la forme de carte 2D.

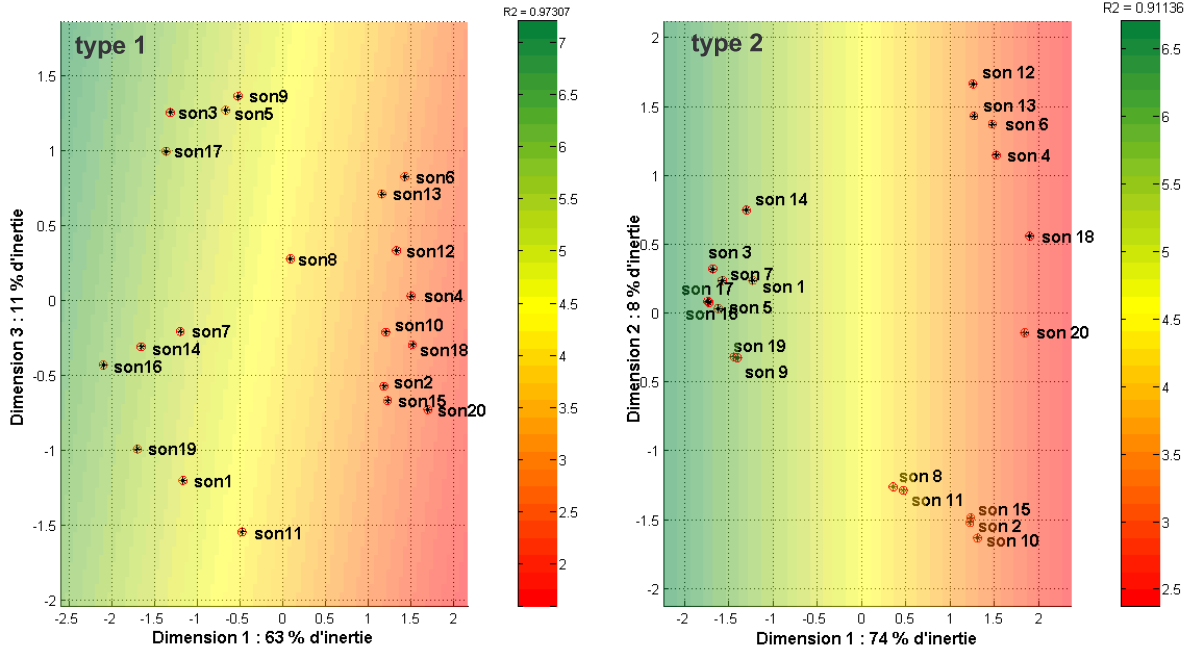


FIG. 12.32 – Cartographie externe des préférences pour les sons de type 1 et 2

Les cartographies externes de préférence pour les sons de type 1 et type 2 sont montrées sur (Fig. 12.32). Le niveau de préférence est marqué par un changement de couleur. Les sons préférés sont présentés en vert, alors que les sons rejetés, en rouge. Le modèle vectoriel décrit très bien les données expérimentales, puisque l'analyse statistique montre que pour les sons de type 1, ce sont 95 % des sons qui sont expliqués par le modèle à $\pm 0.5pts$ et 100 % $\pm 1pt$. Pour les sons de types 2, ce sont 75 % des sons qui sont approximés par le modèle à $\pm 0.5pts$ et 100 % $\pm 1pt$. A l'aide de la cartographie externe on peut observer que pour le type 1, les dimensions les plus importantes pour les préférences sont la dimension 1 (63 % d'inertie) et la dimension 3 (11 % d'inertie), liées au niveau sonore et le tempo des flux. La dimension 2 n'entre pas directement en compte dans l'appréciation des sons de type 1. Pour les sons de type 2, la cartographie externe de préférence ressemble davantage à la représentation MDS sur les deux premières dimensions, liées au niveau sonore et la hauteur. La dimension 1 attire 74 % de l'inertie des préférences, alors que la dimension 2 - 8% de l'inertie.

12.5.2 Discussion des résultats

Globalement, les sons mis dans le même groupe en catégorisation sont évalués d'une manière très proche en termes de préférence. La différence principale concerne les dimensions qui participent au modèle de préférence. Pour les sons de type 2, ce sont les dimensions 1 et 2 qui jouent un rôle important. Pour les sons de type 1, le modèle est mieux approximé avec les dimensions 1 et 3.

Préférences type 1 et type 2 Pour les deux types de sons, les sons les plus appréciés possèdent les mêmes modifications de zone, de même pour les sons moins aimés. Le son le mieux noté pour le type 1 est le son 16, le moins bien noté, le son 20. Pour les sons de type 2, le mieux noté est le son 17, suivi de près de son 16, le moins le son 18" suivi de son 20 (**Fig. 12.33**). Les sons 16, 18, 20 constituent à chacun d'eux un groupe isolé et ce, pour les deux types de sons.

Les sons les moins appréciés Les sons les moins bien notés, son 18 et son 20 ont tous les deux une caractéristique commune. Dans les deux cas ont été augmenté deux zones situé "à la verticale". En effet, cette modification augmente d'une manière significative l'impulsivité du bruit, en rendant la différence de niveau entre deux zones successives en temps très grande. Apparemment, les différences initiales dans le timbre et le niveau des zones des bruits non-modifiés n'ont pas influé le résultat de la modification.

Le groupe de sons 2, 10, 15 est aussi considéré assez désagréable, car très impulsif. Dans ces trois sons, la partie "amortie" est amplifiée. Ceci engendre un niveau global plus élevé par rapport au son sans modification et introduit une impulsivité plus importante. Pour le son 15, l'amplification de la partie "amortie" fait que les zones "attaque" et "amorti" possèdent un niveau relatif similaire, ce qui rend le flux basses fréquences plus rapide, et donc plus impulsif. Il semble que pour les sons 2, 10 et 15 sont un peu irréguliers, au sens où les deux flux, dans les basses et dans les hautes fréquences sont décalés. A l'inverse, les sons bien notés, son 16 et son 14 possèdent une structure rythmique plus déterminé.

Les sons les mieux notés On peut considérer le son 16 comme le mieux noté de tous les sons. Pour les sons de type 2, c'est le son 17 qui est apprécié le plus en moyenne, mais la différence entre lui et le son 16 est minime. Le son 16 possède deux zones atténuées, dans les hautes fréquences ("haut") et dans les basses fréquences ("amortie") (**Fig. 12.34**).

L'atténuation de la zone "haut" est perçue très bien dans tout les sons qui la comportent : son 3, son 16 et son 17. D'un point de vue purement spectral, c'est la zone qui est la plus sensible à l'oreille et qui est censée provoquer le plus un désagrément. Cependant, on peut remarquer que la modification du son 17 du type 2 est mieux apprécié que la même modification pour le type 1. Ceci est certainement dû au fait que initialement, la zone "haut" est plus énergétique pour le type 1 que pour le type 2 et le fait de la diminuer, équilibre les niveaux relatifs entre les zones

Les sons "moyens" Deux sont les sons qui représentent le milieu de l'échelle moyenne de préférence, son 8 et son 11. Les préférences pour ces deux sons sont plutôt "neutres", i.e les sons sont ni particulièrement aimés, ni vraiment rejetés. La modification apportée concerne l'amplification de la partie "attaque" (son 8) où en combinaison avec la diminution de la zone "amortie" (son 11) (**Fig. 12.35**). Cependant, il faut tenir compte que les sons 8 et 11 sont jugés nettement moins bons que le son sans modification, le son 9. Ceci est dû à l'impulsivité plus importante, créée par l'augmentation de la différence de niveau entre les zones adjacentes en temps, "attaque" et "amortie".

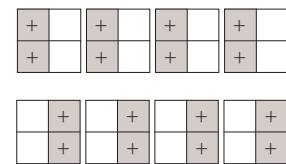


Fig. 12.33 – Schéma des modifications : son 18 (en haut) et son 20 (en bas)

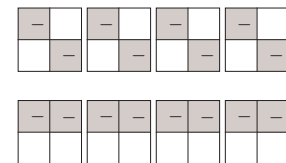


Fig. 12.34 – Schéma des modifications dans le son 16 (en haut) et son 17 (en bas)

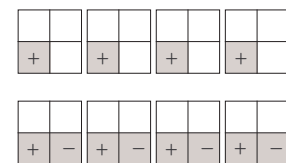


Fig. 12.35 – Schéma des modifications dans le son 8 (en haut) et son 11 (en bas)

Préférences d'un expert moteur Les auditeurs participant dans les tests sont dans leur majorité semi-naïfs, i.e ils connaissent l'univers des bruits moteur, mais ne sont pas de spécialistes en acoustique moteur. Parmi les auditeurs, un seul expert moteur a passé les tests, et la comparaison entre ses réponses peut révéler des différences dans les critères de catégorisation et de préférence.

Le son le mieux noté par l'expert, pour le type 1 et le type 2, est toujours le son 14 (**Fig. 12.36**). Par rapport à la moyenne des réponses des autres auditeurs, le son 14 est globalement bien noté, mais n'est pas le meilleur. Par ailleurs, la meilleure note de l'expert donnée au son 14 est confirmée par le test de catégorisation et par la verbalisation. Une autre grande différence avec les préférences moyennes, est la note plutôt moyenne donnée par l'expert pour le son 16 qui est noté le meilleur pour les auditeurs, et ce pour les deux types de sons. Il semble que le type d'écoute de l'expert est concentré sur les événements temporels en termes de $2N$ et $0.5 N$ moteur et en termes de composantes spectrales, en rapport avec ces événements.

Il semble également que pour l'expert le niveau sonore faible n'est pas un critère déterminant pour la qualité du son. En comparaison, les auditeurs "naïfs" ont nettement préféré les sons moins forts. Ceci est surtout valable pour les sons 16 et 19 du type 1 qui sont notés très bien par les sujets et notés au milieu de l'échelle de préférence par l'expert. Sa justification de choix est l'observation que les sons "sonnent comme si l'harmonique $2N$ du moteur faisait ressortir une fréquence aiguë, en plus d'une autre fréquence et de plus, il n'y a pas de présence de $0.5 N$ moteur". Ce dernier fait, apparemment pas du tout gênant pour les autres sujets, est considéré comme un moins par l'expert. En confirmation, le son le mieux noté par lui, le 14, présente "un claquement $0.5 N$ perceptible mais grave, en combinaison avec une composante $2N$ sourde".

L'expert moteur n'a pas préféré les sons comme 2, 10, 15, 20 pour des raisons d'une impulsivité très marquée et un niveau élevé. En revanche, le son 18, le moins bien noté en moyenne, n'est pas le pire pour l'expert. Ceci révèle que probablement pour l'expert moteur, une modification plutôt temporelle et rythmique, notamment dans la zone "amortie", est beaucoup plus significative qu'une modification spectrale dans les hautes fréquences de l'explosion.

Conclusion sur le test de préférence Le test de préférence montre que, d'une manière générale, les sons plus impulsifs, plus forts et en combinaison avec une composante aiguë très audible, sont moins bien appréciés. Les sons regroupés ensemble par la catégorisation ont une note d'appréciation très proche. Ceci suggère que la préférence est intégrée dans les critères de catégorisation, utilisés par les auditeurs. Il semble aussi que le facteur temporel et rythmique joue également un rôle important, surtout pour les sons de type 1. Ces phénomènes sont liés à la notion d'impulsivité et la formation de plusieurs flux auditifs, en basses et en hautes fréquences, qui entrent en compétition. Enfin, il existe des différences entre l'appréciation d'un expert moteur, qui écoute les bruits en cherchant une équivalence avec les réglages mécaniques du moteur et un auditeur non spécialiste qui juge les bruits moteur en se basant seulement sur leurs propriétés sonores.



Fig. 12.36 – Schéma des modifications : son 14

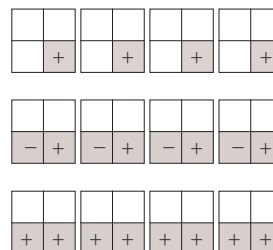


Fig. 12.37 – Schéma des modifications : son 2 (en haut), son 10 (au milieu) et son 15 (en bas)

Conclusion sur les tests perceptifs

Les sons de synthèse sont obtenus à partir de 2 bruits de moteur diesel au ralenti avec des réglages différents. Les deux sons de synthèse respectifs sont à l'origine de deux familles de sons, type 1 et type 2, obtenues à l'aide de modifications des atomes temps-fréquence du signal. Les modifications consistent à amplifier ou atténuer les atomes dans une ou deux zones temps-fréquence, issues d'un découpage du cycle moteur en temps et en fréquence. Un tel découpage permet l'interprétation des bruits moteur au ralenti dans le contexte des flux auditifs et révèle des phénomènes de masquage temps-fréquence. Deux tests perceptifs sont effectués : un test de catégorisation et un test de préférence. Les résultats révèlent l'influence du niveau global des bruits, la hauteur des flux formés par les zones temps-fréquence, ainsi que l'aspect rythmique et impulsif des séquences. La préférence des auditeurs est portée vers des bruits graves dont le niveau est faible. Cette préférence n'est pas entièrement partagée par un expert moteur, dont les choix sont liés à la manifestation sonore d'un bon fonctionnement du moteur.

Perspectives de la thèse : L'hypothèse de la formation des flux auditifs dans la perception des bruits moteur au ralenti est l'idée principale de la troisième partie de la thèse. Cette hypothèse a été confirmée par les tests perceptifs et suggère une nouvelle voie dans l'étude des bruits impulsifs répétitifs. Cette voie est possible grâce à l'algorithme d'analyse-synthèse, développé dans la deuxième partie de la thèse qui permet la manipulation des atomes temps-fréquence du signal de synthèse. L'approche des flux auditifs pour les bruits moteur au ralenti pourra permettre, dans une future recherche, l'association plus fine des phénomènes physiques dans le cycle moteur avec une zone temps-fréquence et la prédiction de l'impact sur la perception des phénomènes physiques dans le cycle moteur.

Apport et perspectives

Cette thèse s'inscrit dans une démarche pluridisciplinaire réunissant les connaissances de trois domaines différents : le bruit automobile, les méthodes de traitement du signal et la psychoacoustique. Un apport essentiel de la thèse est donc cette vision scientifique élargie qui permet une compréhension beaucoup plus complète de la problématique concernée. Par ailleurs, la contribution dans chacun des trois domaines mentionnés ci-dessus comporte des aspects pédagogiques, méthodologiques et conceptuels.

Apport pédagogique La première partie de la thèse est consacrée au large domaine de l'acoustique automobile, en décrivant d'une manière à la fois concise et détaillée les différents aspects du fonctionnement du moteur, ainsi que ses sources primaires et secondaires d'excitation vibro-acoustique. La présentation de l'acoustique automobile est complétée par une description des mécanismes responsables du bruit de roulement et du bruit aérodynamique, ainsi que des processus de transferts aérien et solide des bruits vers l'habitacle. La complexité du bruit automobile est illustrée avec de multiples exemples sonores et graphiques. Un chapitre de la thèse est consacré à un sujet très actuel, celui de la qualité sonore des bruits automobile. La qualité sonore est présentée à l'aide des principaux indicateurs psychoacoustiques et de la méthodologie des tests perceptifs. Elle est influencée également par des facteurs culturels et émotionnels, liés ou pas au contexte automobile. Enfin, une place importante est réservée à l'objet principal de la thèse, les bruits diesel au ralenti. Les particularités des bruits moteur au ralenti sont déterminées par les processus physiques de combustion et d'injection. La nouvelle approche de la thèse met aussi l'accent sur l'importance d'étudier les aspects rythmiques des bruits impulsifs répétitifs afin de comprendre leur perception. D'un point de vue pédagogique, la première partie de la thèse représente une très bonne "entrée en matière" dans la problématique de l'acoustique automobile.

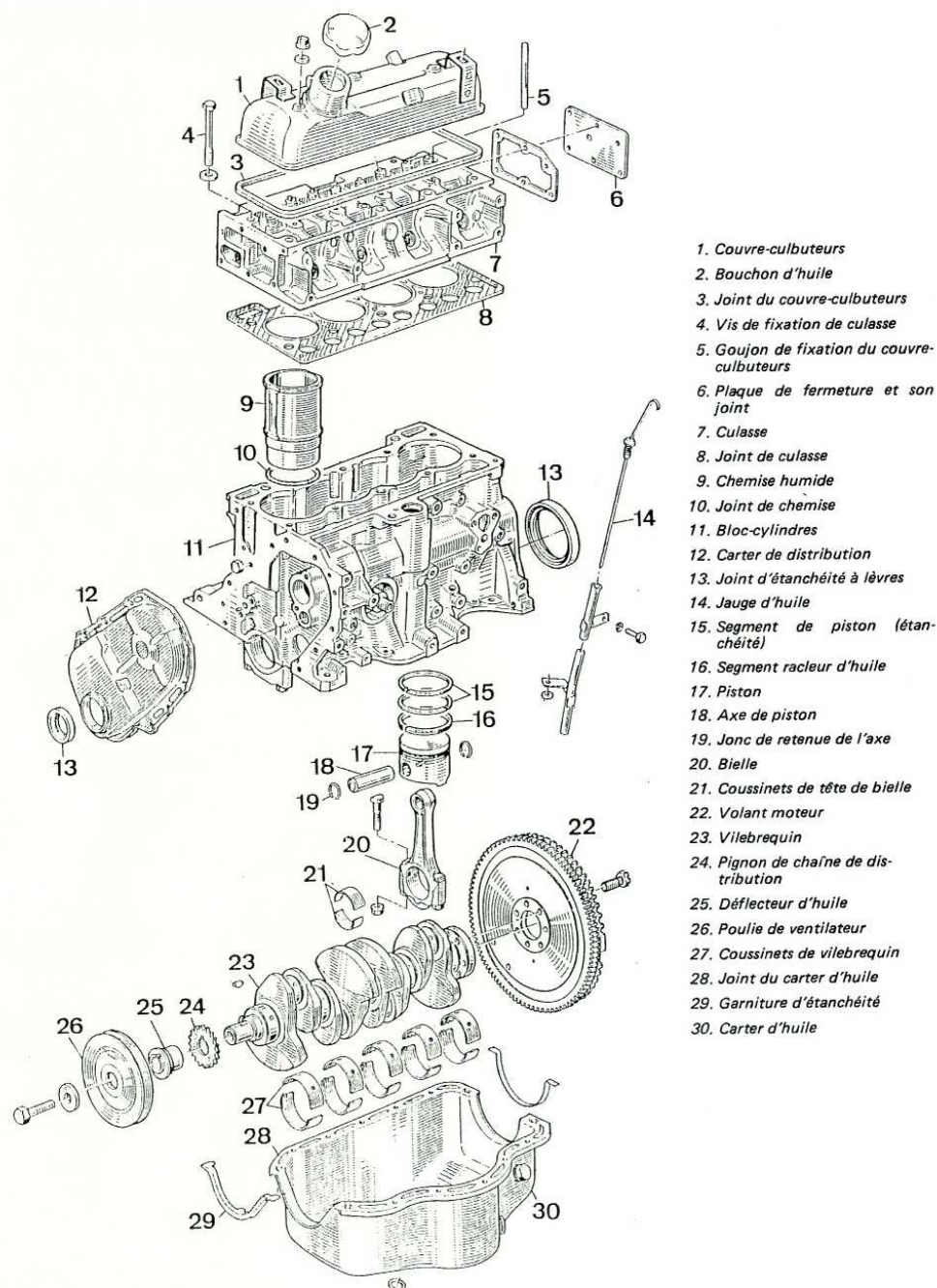
Apport méthodologique L'apport méthodologique principal de la thèse concerne le développement de la nouvelle méthode d'analyse-synthèse, présentée dans la deuxième partie. Cette méthode est spécialement adaptée aux bruits moteur au ralenti, mais peut être appliquée, d'une manière plus générale, à tous les signaux sonores. La nouveauté de la méthode réside dans la combinaison de plusieurs procédés du traitement du signal comme l'analyse temps-fréquence de Gabor, les décompositions parcimonieuses, la réallocation et la synthèse granulaire. Cette approche permet de déduire des formules analytiques pour l'estimation des paramètres des atomes temps-fréquence qui jouent le rôle de nouveaux opérateurs de réallocation. En combinaison avec des critères de sélection des atomes, la méthode d'analyse fournit une décomposition parcimonieuse du signal d'un cycle moteur. La synthèse granulaire construit le bruit de synthèse final par périodisation des atomes. La présentation de la méthode d'analyse-synthèse ne discute pas en détail des problèmes théoriques approfondis comme l'erreur numérique de l'approximation ou la convergence de la réallocation. En revanche, la méthode fournit des signaux de synthèse des bruits au ralenti d'un réalisme satisfaisant laissant conclure sa pertinence. Un avantage de la synthèse granulaire est la possibilité de contrôler en temps et en fréquence le contenu du signal, ce qui a permis de mettre en œuvre une nouvelle interprétation des bruits impulsifs répétitifs dans le contexte des flux auditifs.

Apport conceptuel Un premier apport conceptuel de la thèse se manifeste dans l'approche phénoménologique du bruit moteur adoptée par l'algorithme d'analyse-synthèse. Cette idée considère le bruit moteur non plus d'un point de vue purement spectral, mais plutôt comme une succession d'événements temporels dans le cycle moteur. L'approche phénoménologique est en lien direct avec la deuxième idée innovatrice, exprimée à travers l'interprétation des bruits au ralenti comme une combinaison de plusieurs flux auditifs. La théorie des scènes auditives et les règles de regroupement du Gestalt, discutées dans la troisième partie de la thèse permettent de multiples analogies avec la structure des bruits moteur au ralenti. Un découpage géométrique du cycle moteur en zones temps-fréquence définit une représentation originale des bruits impulsifs répétitifs. Cette approche garde un lien avec des phénomènes physiques et sonores du cycle moteur et permet, grâce à la synthèse granulaire, de modifier indépendamment chaque zone temps-fréquence afin d'étudier son impact perceptif. Les tests perceptifs de catégorisation et de préférence, présentés à la fin de la thèse, appliquent le concept de zones temps-fréquence et mettent en évidence l'existence de flux auditifs dans les basses et les hautes fréquences. Les principales dimensions perceptives des bruits au ralenti déterminées par les tests sont associées au niveau, à la hauteur et à l'aspect rythmique de ces flux auditifs. Ce nouveau concept de zones temps-fréquence a également permis de montrer certains phénomènes de masquage temps-fréquence des bruits moteur au ralenti.

Perspectives Les perspectives de la thèse reposent dans la continuité des deux idées originales développées dans le manuscrit : la méthode d'analyse temps-fréquence et de synthèse granulaire et l'interprétation des bruits moteur au ralenti dans le contexte des flux auditifs. Une future recherche visant l'amélioration de la méthode d'analyse-synthèse peut être concentrée sur l'étude des critères de sélection parcimonieuse des atomes temps-fréquence, des propriétés de convergence et d'optimalité de la décomposition du signal fournie par la méthode, ainsi que sur des problèmes de recouvrement des atomes dans le plan temps-fréquence et leur impact perceptif. Une application sur d'autres signaux sonores, les sons impulsifs en particulier, est également envisageable.

L'idée du découpage schématique du cycle moteur en zones temps-fréquence peut être appliquée sur d'autres signaux à caractère impulsif répétitif, afin d'étudier leurs propriétés psychoacoustiques. Les modifications apportées dans les zones temps-fréquence peuvent concerner non seulement le niveau sonore, mais aussi les autres paramètres des atomes comme la fréquence et le temps d'apparition pour démontrer de divers effets de masquage temps-fréquence. Dans le contexte des bruits moteur, il est intéressant de pouvoir associer plus précisément les zones temps-fréquence avec les réglages du fonctionnement du moteur afin de pouvoir prédire le résultat perceptif sur le bruit. Ces perspectives tracent le lien entre deux mondes : le monde objectif des processus acoustiques et le monde subjectif de la perception sonore.

Les constituants du moteur



Mesures vibroacoustiques

B.1 Introduction

Les mesures vibro-acoustiques sont effectuées au début de la thèse. Leur objectif principal est la familiarisation avec les organes moteur et l'étude de l'ensemble des bruits émis par la voiture. Les enregistrements ont eu lieu le 22-30 juillet 2002 sur les pistes d'essai de Peugeot Citroën à la Ferté Vidame.

B.2 Objectifs des mesures

Les mesures ont pour but de fournir une base de signaux enregistrés simultanément en différents endroits de la voiture. Ces signaux sonores et vibratoires donnent une information plus précise sur les trois contributions principales : le bruit moteur, le bruit de roulement et le bruit aérodynamique. Notre attention s'est portée, en premier, vers le bruit émis par le groupe moto-propulseur directement dans le compartiment moteur. Un intérêt représentent aussi les bruits de roulement et d'écoulement d'air émis à l'extérieur de la voiture, ainsi que leur perception dans l'habitacle. Les mesures acoustiques sont complétées, en quelques endroits choisis, par une étude du comportement vibratoire de la voiture.

B.3 Matériel

Voitures :

- ✓ Peugeot 307
- ✓ Citroën C5

Deux voitures sont choisies : une Peugeot 307, représentant la moyenne gamme, et une Citroën C5, représentant la haute gamme des voitures du Groupe PSA (**Fig. B.1**). Au moment des mesures, les voitures étaient neuves et avec un faible kilométrage (200-500 km). Ces véhicules possèdent de deux systèmes différents de suspension (une suspension hydraulique pour la Citroën et une suspension à ressorts pour la Peugeot), pouvant influencer le comportement acoustique en fonction du type de la route.



FIG. B.1 – Peugeot 307 (à gauche) et Citroën C5 (à droite)

Les profils aérodynamiques différents des deux voitures jouent un rôle important sur le bruit aérodynamique, alors que le volume et la forme de l'habitacle influencent la perception des bruits à l'intérieur.

Moteur :

- ✓ Type EW10J4

Le moteur EW10J4 est un moteur essence à 4 cylindres, avec une injection multipoint (4 soupapes par cylindre) et une cylindrée de 2 l (**Fig. B.2**). Son bruit a été estimé comme plutôt "neutre", sans caractéristiques particulières comme par exemple une rugosité très prononcée. Cette même motorisation a été choisie pour les deux voitures, afin de pouvoir observer à la fois les paramètres caractéristiques propres au moteur seul et les phénomènes de couplage du même moteur avec des structures différentes.

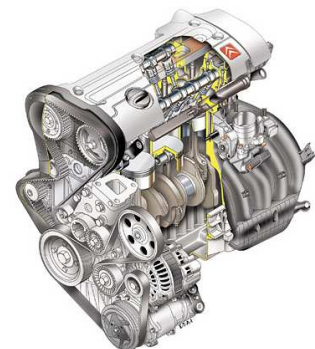


Fig. B.2 – Le moteur EW10J4

Enregistreur numérique :

- ✓ Système d'acquisition De Kerac - 32 voies

La chaîne d'acquisition alimente deux types de voies, microphoniques et accélérométriques. Elle est reliée à un ordinateur portable et possède une partie mémoire, dite frontale, qui permet l'enregistrement et le stockage direct des signaux enregistrés. L'évolution en temps réel de l'enregistrement peut être suivie à l'aide d'un logiciel. L'alimentation de la chaîne est fournie par la batterie de la voiture.

Enregistrements binauraux :

- ✓ Casque microphonique (Fig. B.3)
- ✓ Mannequin d'enregistrement binaural HEAD Acoustics (Fig. B.3)

Le casque microphonique est porté par le conducteur afin de restituer sa perception sonore binaurale. La "tête acoustique", un mannequin enregistrant sur cassette de format DAT et intégrant des facteurs perceptifs du transfert binaural, est directement connecté à la chaîne d'acquisition. Le mannequin est placé sur le siège devant passager.



Capteurs :

- ✓ 19 microphones Ke4
- ✓ 7 accéléromètres ICP

Accessoires :

- ✓ tachymètre
- ✓ anémomètre
- ✓ pistonphone

Fig. B.3 – Casque microphonique (en haut) et tête acoustique (en bas)

B.4 Positionnement des capteurs

Les microphones et les accéléromètres sont placés à l'intérieur de la voiture (dans le compartiment moteur et dans l'habitacle) et à l'extérieur, à côté des roues et du pot d'échappement. Les entrées microphoniques (23 pour l'ensemble de la voiture) sont marquées sur les images ci-dessous avec des cercles, alors que les entrées accélérométriques (7 au total) sont marquées avec des carrés.

B.4.1 Compartiment moteur

Le placement des microphones tient compte des organes moteur mentionnés comme émetteurs de bruit. Le positionnement est aussi influencé par une étude holographique sur le moteur EW10J4. [197].

Etude holographique La méthode de l'holographie acoustique [198], consiste à mesurer le champ acoustique sur une antenne fixe et de le décomposer sur une base d'ondes planes via la transformée de Fourier spatiale (TFS). La rétropropagation sur un plan parallèle permet d'obtenir l'image de la pression phasée reconstituée à l'aide de la TFS inverse. Dans l'étude holographique sur le moteur EW10J4, faite à PSA [197], l'antenne comprend 42 microphones par face, placés sur chacune des 6 faces du moteur. L'étude localise 17 zones avec une émission sonore importante, dont 5 zones sur chacune des faces d'admission et d'échappement, et respectivement 4 et 3 zones sur la face de distribution et la face dessus.

Zones de positionnement Pour nos mesures, dans le compartiment moteur sont mis 13 microphones et 5 accéléromètres. Ces capteurs sont positionnés en tenant compte des zones acoustiques les plus importantes, trouvées par l'étude holographique [197]. Afin d'observer le fonctionnement d'un seul cylindre, plusieurs capteurs sont mis près du troisième cylindre, en comptant de droite à gauche, face au compartiment moteur. Ce cylindre est choisi ainsi afin de capter le moins possible les émissions sonores de la distribution et de la boîte de vitesse. Les capteurs sont mis près de l'injecteur, les soupapes d'admission et d'échappement, le joint culasse. Un bruit rugueux est observé près d'une zone, appelée "cache style centre".

Face dessus

- | | |
|--------------------------------------|---|
| 1. Joint culasse, cylindre 3 | 4. Injecteur, cylindre 3 |
| 2. Soupape d'admission, cylindre 3 | 5. Rampe injecteur, entre les cylindres 2&3 |
| 3. Soupape d'échappement, cylindre 3 | 6. Cache style centre |

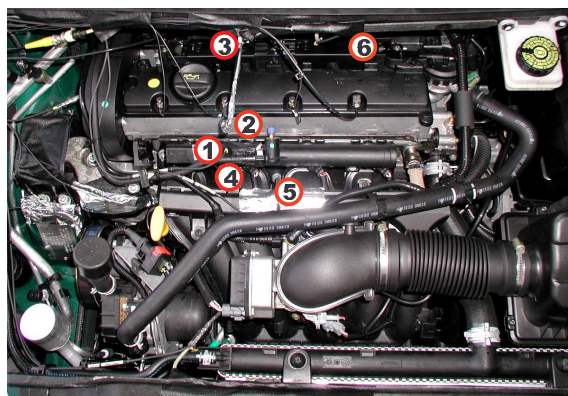
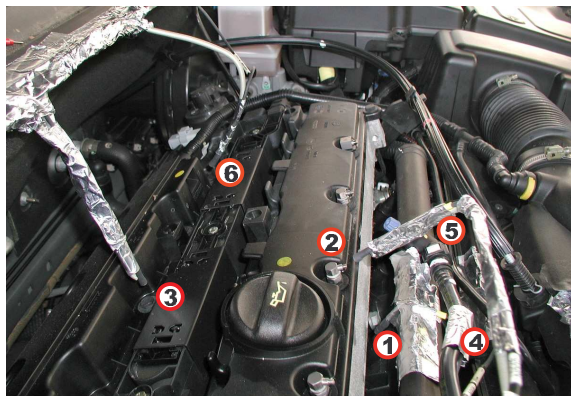


Fig. B.4 – Positions des microphones sur la face dessus du moteur EW10J4. Vue de côté gauche du moteur sur la C5 (à gauche), vue de face du moteur sur la 307 (à droite).

Face distribution**Face admission**

- | | |
|--|--|
| 7. Vilebrequin | 10. Alternateur |
| 8. Entre le galet enrouleur et la courroie | 11. Entre la pompe direction assistée et le boîtier papillon des gaz |
| 9. Entre la poulie et l'arbre à cames | |

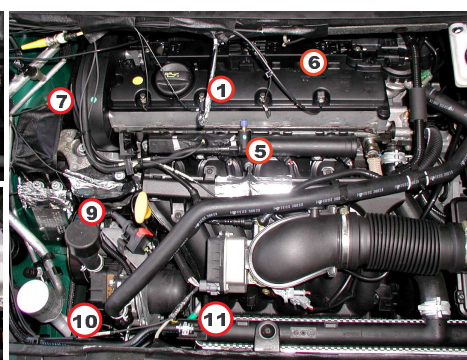
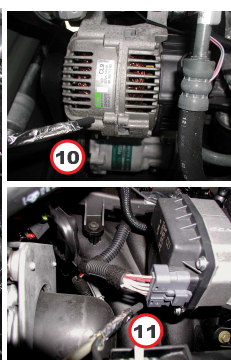
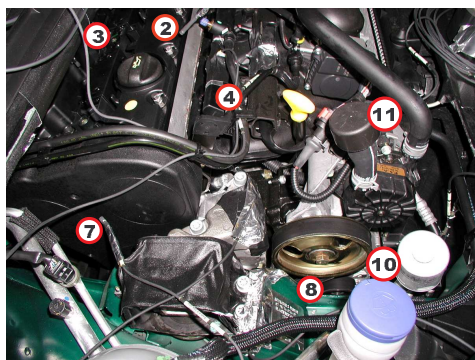


Fig. B.5 – Position des microphones. Côté distribution (à gauche), côté face (à droite), l'alternateur (au milieu, en haut) et la pompe direction assistée (au milieu, en bas).

Face dessous

- | |
|------------------------------|
| 12. Boîte de vitesse |
| 13. Collecteur d'échappement |

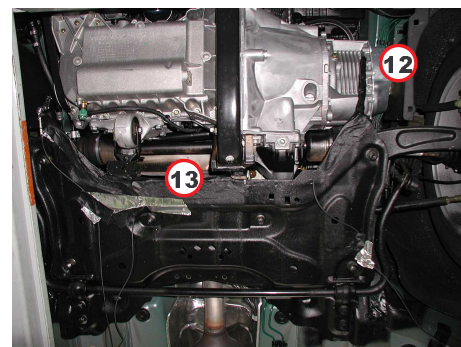
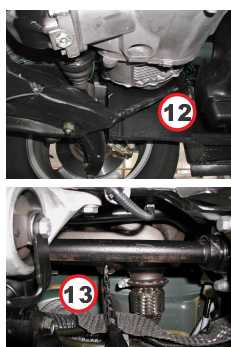


Fig. B.6 – Face dessous (à droite), la boîte de vitesse (à gauche, en haut) et le collecteur d'échappement (à gauche, en bas).

Accéléromètres

1. Cylindre 1, sur la visse moteur
2. Cylindre 2, sur la visse moteur
3. Cylindre 3, sur la visse moteur
4. Cylindre 4, sur la visse moteur
5. Cale moteur

Les accéléromètres sont placés sur les visse moteur. Chaque accéléromètre capte principalement les vibrations liées au cycle moteur d'un seul cylindre. Ceci est fait dans l'objectif de pouvoir corrélérer les cycles individuels des cylindres. L'accéléromètre sur la cale droite du moteur mesure le couplage vibratoire du moteur avec la structure de la voiture (Fig.B.7).

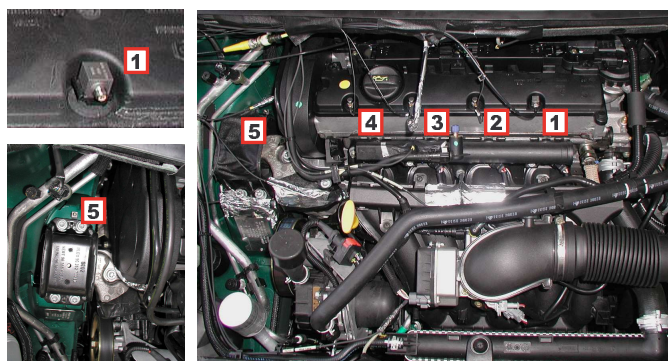


Fig. B.7 – Position des accéléromètres. Face dessus, sur les visse moteur (à droite), une visse (à gauche, en haut), sur la cale moteur (à gauche, en bas).

B.4.2 Roues

14. roue droite avant, en amont
15. roue droite avant, en aval
16. roue droite arrière, en amont
17. roue droite arrière, en aval

Les roues avant et arrière droites, munies de pneus Pirelli 205/55 R16, sont équipées avec 4 microphones portant des bonnettes anti-vent. Les microphones, situés en amont et en aval des deux roues droites, sont à une distance de 10 cm du sol. Les roues droites sont choisies pour limiter les contributions de la boîte de vitesse, située près de la roue avant gauche) et la sortie d'échappement, située près de la roue arrière gauche).



Fig. B.8 – Position des microphones sur les roues avant et arrière droites

B.4.3 Echappement

18. Sortie échappement, avec une bonnette anti-vent (Fig. B.9).

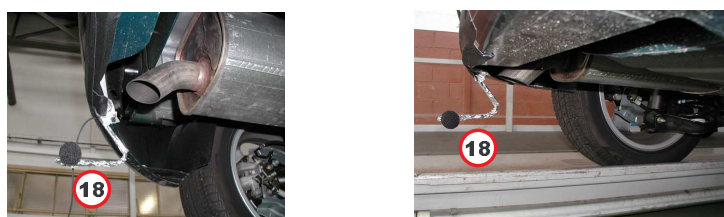


FIG. B.9 – Sortie d'échappement : Peugeot 307 (à gauche), Citroën C5 (à droite)

B.4.4 Habitacle

Dans l'habitacle sont répartis 5 microphones (Fig. B.10) et 2 accéléromètres (Fig.B.11)

19. Tête artificielle HEAD Acoustics sur la place devant passager, voie microphonique stereo droite
20. Tête artificielle HEAD Acoustics sur la place devant passager, voie microphonique stereo gauche
21. Casque microphonique BHM dans les oreilles du conducteur, voie microphonique stereo droite
22. Casque microphonique BHM dans les oreilles du conducteur, voie microphonique stereo gauche
23. Habitacle, derrière le conducteur, microphone près de l'oreille droite

6. Glissière gauche du siège du conducteur
7. Plancher de l'habitacle, derrière le siège du conducteur

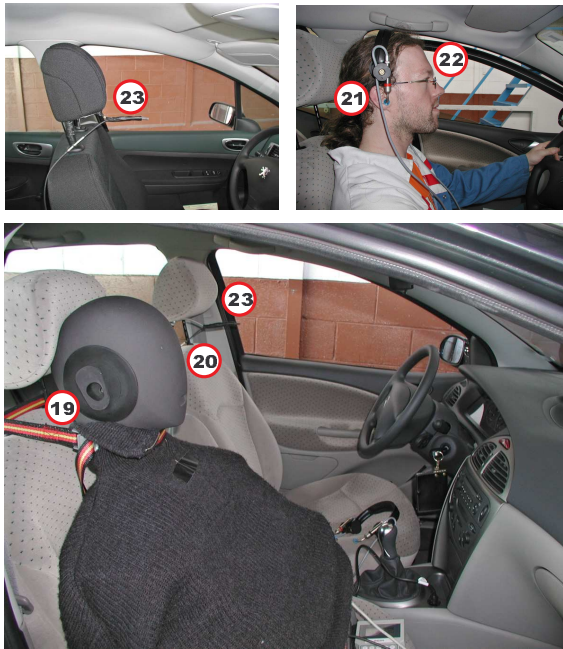


Fig. B.10 – Microphones dans l'habitacle. Tête acoustique (en bas), casque microphonique (en haut, à droite), micro habitacle (en haut, à gauche)

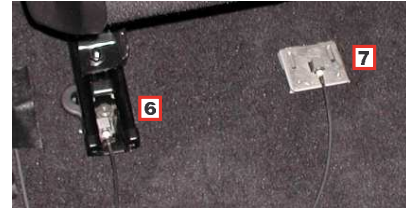


Fig. B.11 – Accéléromètres dans l'habitacle sur la glissière du siège conducteur et le plancher. Vue d'ensemble (en bas), zoom sur les capteurs (en haut)

B.5 Description des mesures

Les mesures sont effectuées dans deux situations de conduite, statique et dynamique, i.e respectivement voiture à l'arrêt et voiture en roulement. Deux pistes différentes sont utilisées, une piste à revêtement normal et une piste avec revêtement rugueux, dite « gros grain ». Les situations de conduite sont choisies afin de pouvoir comparer les sons avec d'autres mesures réalisées, par exemple en soufflerie, sur des pistes et des voitures différentes, etc.

B.5.1 Mesures en statique

L'objectif des mesures à l'arrêt est d'obtenir le bruit de moteur seul, à régimes différents, et sans la contribution du bruit de roulement et d'écoulement d'air. Les enregistrements sont faits en plein air, dans un endroit très calme et avec un capot fermé du compartiment moteur. Les situations sont

- ⇒ Ralenti, environ 750 tr/min
- ⇒ Régimes stabilisés, de 1500 à 6000 tr/min, par pas de 500 tr/min.
- ⇒ Montées contrôlées, de 2000 à 6000 tr/min en augmentant linéairement le régime avec 200 tr/min par seconde.
- ⇒ Montées très rapide, i.e des « coups de pédale » sur l'accélérateur.

Les mesures en statique offrent la possibilité d'étudier les bruits du moteur seul sans avoir l'influence des autres contributions. Toutefois, à l'exception du ralenti, ces cas ne représentent pas une réelle situation de conduite. Il ne prennent pas non plus en compte la charge du moteur et l'inertie de la voiture.

B.5.2 Mesures dynamiques

Les mesures de la voiture en roulement donnent l'ensemble des contributions acoustiques, le bruit moteur et les bruits de roulement et d'écoulement. Les situations de conduite concernent les régimes stabilisés (à un régime moteur et à un rapport de vitesse donnés), les vitesses stabilisées sur deux pistes, la pleine admission, le moteur coupé, le point mort et le changement des rapports de vitesse.

Régimes stabilisés

Ces mesures ont pour objectif de capter les bruits du moteur seul, en tenant compte des paramètres comme le rapport de la boîte de vitesse, la charge, l'inertie du véhicule, etc. Ces paramètres ne sont pas présents dans le cas de mesures en statique. Le tableau (**Fig. B.12**) montre les situations enregistrées, sur une piste "normale", en fonction du rapport de la boîte de vitesse et pour des régimes du moteur fixes. Les vitesses moyennes observées leur correspondant sont indiquées en gras.

regime rapport	2000 tr/min	3000 tr/min	3500 tr/min	4000 tr/min	5000 tr/min
2-ème	30 km/h	47 km/h	52 km/h	60 km/h	72 km/h
3-ème	47 km/h	68 km/h	78 km/h	88 km/h	110 km/h
4-ème	60 km/h	90 km/h	102 km/h	118 km/h	
5-ème	72 km/h	105 km/h	122 km/h		

FIG. B.12 – Tableau récapitulatif des régimes stabilisés mesurés

Vitesses stabilisée

Le bruit de roulement dépend plutôt de la vitesse du véhicule que du régime moteur. Pour les deux voitures, le bruit de roulement est mesuré sur deux pistes différentes, à 3 vitesses constantes.

- ⇒ Vitesses stabilisées - 50, 90 et 140 km/h
- ⇒ Deux pistes - "normale" et "gros grain"

Le revêtement de la piste "gros grain" provoque une excitation vibroacoustique plus importante. Ceci fait ressortir davantage le bruit de roulement, en comparaison avec une route à revêtement normal.

Pleine admission

C'est une situation de "pleine charge" où la pédale d'accélération est appuyée à fond et le papillon des gaz est ouvert à son maximum. Dans cette condition la structure du moteur subit une excitation qui exprime le mieux le bruit caractéristique d'un moteur.

- ⇒ Pleine admission pour un rapport de vitesse 2, 3 et 4, suivie de rétrogradation

Contact coupé

Les mesures du bruit aérodynamique seul nécessitent de limiter les contributions à la fois du bruit de roulement et le bruit moteur. Le bruit de roulement est minimisé en roulant sur une piste lisse. Le bruit moteur est limité en accélérant la voiture jusqu'à une vitesse initiale de 140 km/h, en coupant ensuite le moteur et en laissant la voiture avancer de son inertie. La vitesse choisie est de 140 km/h afin de pouvoir comparer les enregistrements avec des mesures faites en soufflerie. Deux situations sont étudiées :

- ⇒ Contact coupé, papillon de gaz ouvert
- ⇒ Contact coupé, papillon de gaz fermé

"Point mort"

Une situation similaire à la précédente, où la voiture est emmenée à une vitesse de 140 km/h. Ensuite la pédale d'accélération est lâchée, et la voiture est laissée en un "point mort", i.e à sa propre inertie.

- ⇒ Lâchée de pédale - piste "normale"
- ⇒ Lâchée de pédale - piste "gros grain"

Changement de vitesse

Le changement du rapport de la boîte de vitesse possède un bruit particulier. Dans cet essai, le rapport de vitesse est augmenté à chaque fois quand le régime moteur atteint 3000 tr/min.

- ⇒ Changement de vitesse de 2-ème jusqu'à 5-ème à 3000 tr/min

B.6 Traitement des données

Après l'enregistrement, les données sont stockées dans la mémoire de la chaîne d'acquisition, dans des fichiers de format compressé avec l'extension *.mux*. Ces fichiers contiennent l'information de tous les capteurs (voies microphoniques et voies accélérométriques) pour chaque séquence, i.e situation de conduite. Le processus de traitement consiste à extraire les données à l'aide d'un logiciel, voie par voie et séquence par séquence, de créer des fichiers intermédiaires *.txt* et de les exporter ensuite comme signaux en format *.wav* sous MATLAB.

Liste des exemples sonores

- Piste 1 :** Signal accélérométrique d'un moteur à essence au ralenti, enregistré sur la visse moteur du 1-er cylindre (côté transmission). Voir Annexe B.
- Piste 2 :** Bruit moteur à 3000 tr/min, enregistré à proximité de la soupape d'échappement du 3-ème cylindre d'un moteur à essence. Voir Annexe B.
- Piste 3 :** Bruit moteur avec et sans bourdonnement. La graphique illustre le spectre d'une sélection de 2 s de chaque type de bruit, à un même régime, extraite à partir de la longue séquence. Source : Glossaire Acoustique Renault, Piste 1.
- Piste 4 :** Bruit de pilonnement moteur à 4200 tr/min. La graphique illustre le spectre d'une sélection de 3 s extraite à partir de la longue séquence [37-40 s]. Source : Glossaire Acoustique Renault, Piste 68.
- Piste 5 :** Rudesse moteur. Source séquence : Glossaire Acoustique Renault, Piste 74
- Piste 6 :** Bruit des soupapes et l'injecteur, enregistré sur un moteur à essence au ralenti, à proximité de la soupape d'admission et préalablement filtré à l'aide d'un filtre passe-haut. Voir Annexe B.
- Piste 7 :** Bruit d'un moteur à essence en régime stabilisé à 4000 tr/min, rapport 2-nde, enregistré près de la soupape d'admission du cylindre 3. Voir Annexe B.
- Piste 8 :** Bruit d'un moteur à essence en régime stabilisé à 4000 tr/min, rapport 2-nde, enregistré près du collecteur d'échappement. Voir Annexe B.
- Piste 9 :** Bruit d'un moteur à essence, en régime stabilisé à 5000 tr/min, rapport 2-nde, enregistré sur le côté de la transmission moteur, entre la courroie et le galet tendeur. Voir Annexe B.
- Piste 10 :** Bruit de sirènement de la boîte vitesse. Piste sonore extrait de la séquence Piste 75, Glossaire Acoustique Renault
- Piste 11 :** Bruit d'un moteur à essence en régime stabilisé à 4000 tr/min, rapport 2-nde, enregistré près de l'alternateur. Voir Annexe B.
- Piste 12 :** Bruit d'un moteur à essence en régime stabilisé à 5000 tr/min, rapport 2-nde, enregistré près de l'alternateur. Voir Annexe B.
- Piste 13 :** Grogement de la pompe direction assistée. Séquence [2-4 s], extraite de la Piste 41. Source : Glossaire Acoustique Renault
- Piste 14 :** Sirènement de la pompe direction assistée. Séquence, extraite de la Piste 43. Source : Glossaire Acoustique Renault
- Piste 15 :** Bruit de roulement sur piste à revêtement standard, à 50 km/h, rapport 3-ème, enregistré en amont de la roue arrière. Voir Annexe B.
- Piste 16 :** Bruit de roulement sur piste à revêtement standard, à 50 km/h, rapport 3-ème, enregistré dans l'habitacle à l'aide d'une tête acoustique (le signal mono gauche). Voir Annexe B.
- Piste 17 :** Bruit de roulement sur piste à revêtement rugueux ("gros grain"), à 50 km/h, rapport 3-ème, enregistré en amont de la roue arrière. Voir Annexe B.
- Piste 18 :** Bruit de roulement sur piste à revêtement rugueux ("gros grain"), à 50 km/h, rapport 3-ème, enregistré dans l'habitacle à l'aide d'une tête acoustique (le signal mono gauche). Voir Annexe B.
- Piste 19 :** Bruit de roulement sur piste à revêtement standard, à 90 km/h, rapport 4-ème, enregistré en amont de la roue arrière. Voir Annexe B.
- Piste 20 :** Bruit de roulement sur piste à revêtement standard, à 90 km/h, rapport 4-ème, enregistré dans l'habitacle à l'aide d'une tête acoustique (le signal mono gauche). Voir Annexe B.
- Piste 21 :** Bruit de roulement sur piste à revêtement rugueux ("gros grain"), à 90 km/h, rapport 4-ème, enregistré en amont de la roue arrière. Voir Annexe B.
- Piste 22 :** Bruit de roulement sur piste à revêtement rugueux ("gros grain"), à 90 km/h, rapport 4-ème, enregistré dans l'habitacle à l'aide d'une tête acoustique (le signal mono gauche). Voir Annexe B.
- Piste 23 :** Bruit aéroacoustique. Bruit total [2-6 s], bruit aéro isolé [9-13 s], bruit total [16-20 s], sans bruit aéro [23-27 s]. Séquence, extraite de la Piste 4. Source : Glossaire Acoustique Renault

- Piste 24** : Bruit aéroacoustique de 3 voitures à profil différent. Séquence, extraite de la Piste 5. Source : Glossaire Acoustique Renault
- Piste 25** : Avec et sans moteur, papillon des gaz ouvert, enregistré sur piste lisse. Accéléromètre de la cale droite. Voir Annexe B.
- Piste 26** : Avec et sans moteur, papillon des gaz ouvert, enregistré sur piste lisse. Accéléromètre de la glissière du siège conducteur. Voir Annexe B.
- Piste 27** : Avec et sans moteur, papillon des gaz ouvert, enregistré sur piste lisse. Accéléromètre du plancher. Voir Annexe B.
- Piste 28** : Transfert solidien vers l'habitacle, revêtement standard. Accéléromètre de la cale droite, rapport vitesse 3eme, 90 km/h. Voir Annexe B.
- Piste 29** : Transfert solidien vers l'habitacle, revêtement standard. Accéléromètre de la glissière du siège conducteur, rapport vitesse 3eme, 90 km/h. Voir Annexe B.
- Piste 30** : Transfert solidien vers l'habitacle, revêtement standard. Accéléromètre du plancher, rapport vitesse 3eme, 90 km/h. Voir Annexe B.
- Piste 31** : Transfert solidien vers l'habitacle, revêtement rugueux ("gros grain"). Accéléromètre de la cale droite, rapport vitesse 3eme, 90 km/h. Voir Annexe B.
- Piste 32** : Transfert solidien vers l'habitacle, revêtement rugueux ("gros grain"). Accéléromètre de la glissière du siège conducteur, rapport vitesse 3eme, 90 km/h. Voir Annexe B.
- Piste 33** : Transfert solidien vers l'habitacle, revêtement rugueux ("gros grain"). Accéléromètre du plancher, rapport vitesse 3eme, 90 km/h. Voir Annexe B.
- Piste 34** : Transfert aérien vers l'habitacle. Avec et sans moteur, papillon des gaz ouvert, sur revêtement lisse. Enregistré près du joint culasse de cylindre 3. Voir Annexe B.
- Piste 35** : Transfert aérien vers l'habitacle. Avec et sans moteur, papillon des gaz ouvert, sur revêtement lisse. Enregistré près de l'oreille droite du conducteur. Voir Annexe B.
- Piste 36** : Transfert aérien vers l'habitacle. Avec et sans moteur, papillon des gaz ouvert, sur revêtement lisse. Enregistré près de la roue arrière amont. Voir Annexe B.
- Piste 37** : Ralenti diesel avec combustion. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën.
- Piste 38** : Ralenti diesel sans combustion. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën.
- Piste 39** : Ralenti diesel avec injection pilote. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën.
- Piste 40** : Ralenti diesel sans injection pilote. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën.
- Piste 41** : Moteur diesel à froid et à chaud. Source : Glossaire Acoustique Renault, Piste 34.
- Piste 42** : Moteur diesel à froid et à chaud. Source : Glossaire Acoustique Renault, Piste 35.
- Piste 43** : Configuration rythmique 1234 d'un cycle moteur diesel au ralenti. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën.
- Piste 44** : Configuration rythmique 1111 d'un cycle moteur diesel au ralenti.
- Piste 45** : Configuration rythmique 2222 d'un cycle moteur diesel au ralenti.
- Piste 46** : Configuration rythmique 1112 d'un cycle moteur diesel au ralenti.
- Piste 47** : Configuration rythmique 2221 d'un cycle moteur diesel au ralenti.
- Piste 48** : Configuration rythmique 1221 d'un cycle moteur diesel au ralenti.
- Piste 49** : Configuration rythmique 2211 d'un cycle moteur diesel au ralenti.
- Piste 50** : Configuration rythmique 1212 d'un cycle moteur diesel au ralenti.
- Piste 51** : Diesel au ralenti 750 trm, signal original. Sans injection pilote. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën.
- Piste 52** : Bruit de synthèse avec un paramètre de la fenêtre d'analyse $\sigma = 20$.
- Piste 53** : Bruit de synthèse avec un paramètre de la fenêtre d'analyse $\sigma = 25$.
- Piste 54** : Bruit de synthèse avec un paramètre de la fenêtre d'analyse $\sigma = 33$.
- Piste 55** : Bruit de synthèse avec un paramètre de la fenêtre d'analyse $\sigma = 35$.

- Piste 56 :** Bruit de synthèse avec un paramètre de la fenêtre d'analyse $\sigma = 40$.
- Piste 57 :** Bruit de synthèse avec un paramètre de la fenêtre d'analyse $\sigma = 50$.
- Piste 58 :** Diesel au ralenti 750 trm, signal original. Sans injection pilote. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën.
- Piste 59 :** Bruit de synthèse avec 585 maxima locaux.
- Piste 60 :** Bruit de synthèse final après sélection avec 220 atomes.
- Piste 61 :** Bruit de synthèse avec 324 atomes, seuil de produit scalaire = 0.5
- Piste 62 :** Bruit de synthèse avec 244 atomes, seuil de produit scalaire = 0.1
- Piste 63 :** Bruit de synthèse avec 124 atomes, seuil de produit scalaire = 0.001
- Piste 64 :** Bruit de synthèse avec 82 atomes, seuil de produit scalaire = 0.00001
- Piste 65 :** Bruit de synthèse sur 1 cycle moteur
- Piste 66 :** Bruit de synthèse sur 2 cycles moteur
- Piste 67 :** Bruit de synthèse sur 3 cycles moteur
- Piste 68 :** Bruit de synthèse sur 4 cycles moteur
- Piste 69 :** Bruit de synthèse sur 10 cycles moteur
- Piste 70 :** Permutation aléatoire des cordonnées des atomes sur 10% du cycle moteur.
- Piste 71 :** Permutation aléatoire des cordonnées des atomes sur 30% du cycle moteur.
- Piste 72 :** Bruit moteur à 3000 tr/min enregistré près de la soupape d'échappement. Voir Annexe B.
- Piste 73 :** Bruit de synthèse à 3000 tr/min.
- Piste 74 :** Bruit dans l'habitacle, enregistré à l'aide d'une tête acoustique (le signal mono gauche). Rapport vitesse 4, 90 km/h. Voir Annexe B.
- Piste 75 :** Bruit de synthèse dans l'habitacle.
- Piste 76 :** Ségrégation de flux grave et aigu dans une sonate de Telemann. Exemple 6 du CD de démonstration de "Auditory scene analysis" de Bregman.
- Piste 77 :** Ségrégation de flux grave et aigu dans un trille. Exemple 3 du CD de démonstration de "Auditory scene analysis" de Bregman.
- Piste 78 :** Ségrégation de bruits bande étroite. Exemple 14 du CD de démonstration de "Auditory scene analysis" de Bregman.
- Piste 79 :** Formation de résidu grave et aigu. Exemple 34 du CD de démonstration de "Auditory scene analysis" de Bregman.
- Piste 80 :** Formation de 3 flux : grave, moyen et aigu. Exemple 35 du CD de démonstration de "Auditory scene analysis" de Bregman.
- Piste 81 :** Perception des harmoniques en continue. "Au clair de lune", où chaque note est obtenue à partir de l'enregistrement du bruit moteur en statique (1500-5000 trm). Voir Annexe B.
- Piste 82 :** Premier quart d'un cycle moteur au ralenti. Sans injection pilote. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën.
- Piste 83 :** Deuxième quart d'un cycle moteur au ralenti.
- Piste 84 :** Troisième quart d'un cycle moteur au ralenti.
- Piste 85 :** Quatrième quart d'un cycle moteur au ralenti.
- Piste 86 :** Rythme de samba, pandeiro. Source Internet.
- Piste 87 :** Ralenti moteur sous-échantillonné.
- Piste 88 :** Bruit de synthèse au ralenti.
- Piste 89 :** Bruit de synthèse au ralenti avec soustraction de la zone 2-5 kHz.
- Piste 90 :** Bruit de synthèse au ralenti.
- Piste 91 :** Bruit de synthèse au ralenti avec soustraction de la zone 1-4 kHz.
- Piste 92 :** Bruit de synthèse au ralenti.
- Piste 93 :** Bruit de synthèse au ralenti avec soustraction de 3 atomes dans la zone 43-45 ms en temps et 1-7 kHz en fréquence.
- Piste 94 :** Bruit de synthèse au ralenti avec soustraction de 3 atomes dans la zone 20-35 ms en temps et 2-4 kHz en fréquence.

- Piste 95 :** Bruit de synthèse au ralenti avec soustraction de la zone 0-20 ms en temps et 0-12 kHz en fréquence.
- Piste 96 :** Bruit de synthèse au ralenti avec soustraction de la zone 40-60 ms en temps et 0-12 kHz en fréquence.
- Piste 97 :** Bruit de synthèse au ralenti avec soustraction de la zone 80-100 ms en temps et 0-12 kHz en fréquence.
- Piste 98 :** Bruit de synthèse au ralenti avec soustraction de la zone 120-140 ms en temps et 0-12 kHz en fréquence.
- Piste 99 :** Bruit de synthèse au ralenti avec soustraction de la zone 0-20 ms en temps et 1-5 kHz en fréquence.
- Piste 100 :** Bruit de synthèse au ralenti avec soustraction de la zone 40-60 ms en temps et 1-5 kHz en fréquence.
- Piste 101 :** Bruit de synthèse au ralenti avec soustraction de la zone 80-100 ms en temps et 1-5 kHz en fréquence.
- Piste 102 :** Bruit de synthèse au ralenti avec soustraction de la zone 120-140 ms en temps et 1-5 kHz en fréquence.
- Piste 103 :** Bruit de synthèse au ralenti avec soustraction de la zone 140-160 ms en temps et 1-5 kHz en fréquence.
- Piste 104 :** Bruit de synthèse avec soustraction des zones 10-30ms, 50-70ms, 90-110ms, et 130-150ms en temps et 1-5kHz en fréquence.
- Piste 105 :** Bruit de synthèse avec soustraction des zones 20-40ms, 60-80ms, 100-120ms, et 140-160ms en temps et 1-5kHz en fréquence.
- Piste 106 :** Bruit de synthèse avec soustraction alternative des zones 0-10ms en temps et 0-4kHz en fréquence et 40-50ms en temps et 4-12kHz en fréquence.
- Piste 107 :** Diesel au ralenti à 750 trm, signal original. Sans injection pilote. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën. Son 1 original.
- Piste 108 :** Bruit de synthèse du signal original de la Piste 107. Paramètre de la fenêtre d'analyse $\sigma = 25$, synthèse sur 1 cycle.
- Piste 109 :** Diesel au ralenti à 750 trm, signal original. Avec injection pilote. Moteur DV4 au banc. Enregistrement par l'équipe ACV, PSA Peugeot Citroën. Son 2 original.
- Piste 110 :** Bruit de synthèse du signal original de la Piste 109. Paramètre de la fenêtre d'analyse $\sigma = 25$, synthèse sur 1 cycle.
- Piste 111 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "attaque" du cylindre 1, périodisé sur 15 cycles.
- Piste 112 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "attaque" du cylindre 2, périodisé sur 15 cycles.
- Piste 113 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "attaque" du cylindre 3, périodisé sur 15 cycles.
- Piste 114 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "attaque" du cylindre 4, périodisé sur 15 cycles.
- Piste 115 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "amortie" du cylindre 1, périodisé sur 15 cycles.
- Piste 116 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "amortie" du cylindre 2, périodisé sur 15 cycles.
- Piste 117 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "amortie" du cylindre 3, périodisé sur 15 cycles.
- Piste 118 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "amortie" du cylindre 4, périodisé sur 15 cycles.
- Piste 119 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "haut" du cylindre 1, périodisé sur 15 cycles.
- Piste 120 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "haut" du cylindre 2, périodisé sur 15 cycles.

- Piste 121 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "haut" du cylindre 3, périodisé sur 15 cycles.
- Piste 122 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "haut" du cylindre 4, périodisé sur 15 cycles.
- Piste 123 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "résidu" du cylindre 1, périodisé sur 15 cycles.
- Piste 124 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "résidu" du cylindre 2, périodisé sur 15 cycles.
- Piste 125 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "résidu" du cylindre 3, périodisé sur 15 cycles.
- Piste 126 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 107. Zone "résidu" du cylindre 4, périodisé sur 15 cycles.
- Piste 127 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "attaque" du cylindre 1, périodisé sur 15 cycles.
- Piste 128 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "attaque" du cylindre 2, périodisé sur 15 cycles.
- Piste 129 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "attaque" du cylindre 3, périodisé sur 15 cycles.
- Piste 130 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "attaque" du cylindre 4, périodisé sur 15 cycles.
- Piste 131 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "amortie" du cylindre 1, périodisé sur 15 cycles.
- Piste 132 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "amortie" du cylindre 2, périodisé sur 15 cycles.
- Piste 133 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "amortie" du cylindre 3, périodisé sur 15 cycles.
- Piste 134 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "amortie" du cylindre 4, périodisé sur 15 cycles.
- Piste 135 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "haut" du cylindre 1, périodisé sur 15 cycles.
- Piste 136 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "haut" du cylindre 2, périodisé sur 15 cycles.
- Piste 137 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "haut" du cylindre 3, périodisé sur 15 cycles.
- Piste 138 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "haut" du cylindre 4, périodisé sur 15 cycles.
- Piste 139 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "résidu" du cylindre 1, périodisé sur 15 cycles.
- Piste 140 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "résidu" du cylindre 2, périodisé sur 15 cycles.
- Piste 141 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "résidu" du cylindre 3, périodisé sur 15 cycles.
- Piste 142 :** Extraction de zone temps-fréquence à partir du signal original de la Piste 109. Zone "résidu" du cylindre 4, périodisé sur 15 cycles.
- Piste 143 :** Son 1, son de synthèse de Type 1,
- Piste 144 :** Son 2, son de synthèse de Type 1,
- Piste 145 :** Son 3, son de synthèse de Type 1,
- Piste 146 :** Son 4, son de synthèse de Type 1,
- Piste 147 :** Son 5, son de synthèse de Type 1,
- Piste 148 :** Son 6, son de synthèse de Type 1,
- Piste 149 :** Son 7, son de synthèse de Type 1,
- Piste 150 :** Son 8, son de synthèse de Type 1,

Piste 151 : Son 9, son de synthèse de Type 1, son de synthèse sans modification, identique à celui de la piste 107

Piste 152 : Son 10, son de synthèse de Type 1,

Piste 153 : Son 11, son de synthèse de Type 1,

Piste 154 : Son 12, son de synthèse de Type 1,

Piste 155 : Son 13, son de synthèse de Type 1,

Piste 156 : Son 14, son de synthèse de Type 1,

Piste 157 : Son 15, son de synthèse de Type 1,

Piste 158 : Son 16, son de synthèse de Type 1,

Piste 159 : Son 17, son de synthèse de Type 1,

Piste 160 : Son 18, son de synthèse de Type 1,

Piste 161 : Son 19, son de synthèse de Type 1,

Piste 162 : Son 20, son de synthèse de Type 1,

Piste 163 : Son 1, son de synthèse de Type 2,

Piste 164 : Son 2, son de synthèse de Type 2,

Piste 165 : Son 3, son de synthèse de Type 2,

Piste 166 : Son 4, son de synthèse de Type 2,

Piste 167 : Son 5, son de synthèse de Type 2,

Piste 168 : Son 6, son de synthèse de Type 2,

Piste 169 : Son 7, son de synthèse de Type 2,

Piste 170 : Son 8, son de synthèse de Type 2,

Piste 171 : Son 9, son de synthèse de Type 2, son de synthèse sans modification, identique à celui de la piste 109

Piste 172 : Son 10, son de synthèse de Type 2,

Piste 173 : Son 11, son de synthèse de Type 2,

Piste 174 : Son 12, son de synthèse de Type 2,

Piste 175 : Son 13, son de synthèse de Type 2,

Piste 176 : Son 14, son de synthèse de Type 2,

Piste 177 : Son 15, son de synthèse de Type 2,

Piste 178 : Son 16, son de synthèse de Type 2,

Piste 179 : Son 17, son de synthèse de Type 2,

Piste 180 : Son 18, son de synthèse de Type 2,

Piste 181 : Son 19, son de synthèse de Type 2,

Piste 182 : Son 20, son de synthèse de Type 2,

Bibliographie

- [1] **Bensaïd B.** « Les carburants routiers en Europe : l'explosion de la demande en gazole ». Note de synthèse : Panorama 2005. *Institut Français du Pétrole*, (2005).
- [2] **Luigi Russolo.** Source Internet : .
- [3] **Citroën.** Source Internet : <http://www.citroen.fr>.
- [4] **Derreumaux B.** « Moteurs à essence », vol. 1. *ETAI*, (1991).
- [5] « Special Acoustique Vibration ». *Bulletin technique Renault*, 3 (1991). Hors série.
- [6] **Cooper J.** « Etude du groupe motopropulseur ». Rapport bibliographique de thèse. Note interne *PSA/DINQ/DRIA/SARA/PEFH*, (2001).
- [7] **Antoni G.** « Apports de l'échantillonnage angulaire et de la cyclostationnarité au diagnostic par analyse vibratoire des moteurs thermiques ». Thèse de doctorat, *Institut National Polytechnique de Grenoble*, (2000).
- [8] **Rossi J.P.** « Classification des moteurs diesel ». Source Internet : <http://perso.orange.fr/jean-pierre.rossi/diesel/diesel.htm>.
- [9] **Glossaire moteur** :. « Injecteur-pompe ». Source Internet : <http://www.auto-innovations.com/site/glossaire/gmoteur-inj.html>.
- [10] **French M., Nowland M., Hedges D. and Greeley D.** « Reducing noise from gasoline engine injectors ». *International Journal of Vehicle Noise and Vibration*, 1 : p.83–96, (2004).
- [11] *Synthèse Acoustique et Vibrations. Glossaire acoustique.* Renault.
- [12] **Barthod M., Hayne B and Tebec J.-L.** « Auditory perception of noise known as rattle in gearboxes ». *Acta Acustica united with Acustica*, 91 : p.180–191, (2005).
- [13] **Holographie acoustique.** Source Internet : <http://www.mesures.com/archives/755solholographie.pdf>.
- [14] **Williams E. G. and Maynard J. D.** « Holographic imaging without the wavelength resolution limit ». *Physics Review Letters*, 45 : p.554–557, (1980).
- [15] **Hyvärinen A. and Oja E.** « independent component analysis : Algorithms and applications ». *Neural Networks*, 13(4-5) : p.411–430, (2000).
- [16] **Li W., Gu F., Ball A. D. and Leung A.Y.T.** « the identification of diesel engine noise sources ». Dans *ISMA 25, Proceedings*, vol. 13-15/09, p. 1277–1283, (2000).
- [17] **FastICA.** Source Internet : <http://www.cis.hut.fi/projects/ica/fastica/>.
- [18] **Norme de bruit de roulement.** Source Internet : <http://www.ile-de-france.equipement.gouv.fr>.
- [19] **Bruit de roulement.** Source Internet : <http://etat.geneve.ch/dt/environnement/accueil.html>.
- [20] **Bruit de roulement.** Source Internet : <http://www.michelintransport.com>.
- [21] **Brosseaud Y. and Anfosso-Lédée F.** « Les revêtements de chaussées limitant le bruit de roulement, exemple de partenariat et de coopération entre l'administration et les entreprises françaises ». Dans *NFRA 2001*, Montreal, (2001).
- [22] **Van Keulen W. and Duskov M.** « Inventory study of basic knowledge on tyre/road noise ». Rapport DWW-2005-022. *Innovatie Programma Geluid (IPG)*, October (2005).
- [23] **Texture des chaussées.** Source Internet : <http://www.arts-et-metiers.net>.
- [24] **O'Boy D. J.** « Tyre/road interaction noise - a summary ». Source Internet : <http://www2.eng.cam.ac.uk/~djo31/TyreNoise.html>.
- [25] **Bruit de roulement.** « Avis adopté par le conseil économique et social au cours de sa séance du 15 avril 1998 ». Source Internet : <http://www.irma.asso.fr/IMG/pdf/98041506.pdf>.
- [26] **Sandberg U. and Edjsmont J.** « Tire/road noise reference book ». *Informex*, (2002).
- [27] **A. Kuijpers and Van Blokland G.** « Tire/road noise models in the last two decades : a critical evaluation ». Dans *Proceedings InterNoise 2001*, Hague, The Netherlands (2001).
- [28] **Schaff K. and Ronnenberger D.** « Noise radiation from rolling tyres - sound amplification by the horn effect ». Dans *Proceedings InterNoise 1982*, San Francisco, California (1982).

- [29] **Vergne S. and Auger J.-M.** « Aéroacoustique automobile : Analyse des mécanismes en jeu ». Note interne *PSA*.
- [30] **Lighthill M. J.** « On the sound generated aerodynamically : I. General theory ». *Proceedings of the Royal Society of London*, 211(A) : p.564–587, (1952).
- [31] **Goncalvès E.** « Elements d'aéroacoustique ». *Institut National Polytechnique de Grenoble*, mars (2005). Source Internet : <http://www.librecours.org/documents/70/7005.pdf>.
- [32] **Curle N.** « The influence of solid boundary upon aerodynamic sound ». *Proceedings of the Royal Society London*, 231(A) : p.505–514, (1955).
- [33] **Powell A.** « Theory of vortex sound ». *Journal of the Acoustical Society of America*, 36 : p.177–195, (1964).
- [34] **Aérodynamique des véhicules.** Source Internet : <http://membres.lycos.fr/arnoral/principes4.htm>.
- [35] **Chemins de transfert.** Source Internet : <http://www.head-acoustics.de>.
- [36] **Zwicker E. and Fastl H.** « Psychoacoustics : Facts and Models ». *Springer*, (1990).
- [37] **Van Noorden L.** « Temporal coherence in the perception of tone sequences ». Thèse de doctorat, *Eindhoven University of Technology*, Netherlands, (1975).
- [38] **Bregman A. S.** « Auditory Scene Analysis ». *Cambridge, MA : MIT Press*, (1990).
- [39] **McAdams S. and Bigand E.** « Penser les sons. Psychologie cognitive de l'audition », Chapitre : Introduction à la cognition auditive. *Presses Universitaires de France*, (1994).
- [40] **Gibson J.J.** « The Ecological Approach to Visual Perception ». *Houghton Mifflin*, Boston, (1979).
- [41] **Roussarie V.** « Analyse perceptive des structures vibrantes ». Thèse de doctorat, *Université du Maine*, (1999).
- [42] **Guyot F.** « Etude de la perception sonore en termes de reconnaissance et d'appréciation qualitative : une approche par la catégorisation ». Thèse de doctorat, *Université du Maine*, (1996).
- [43] **Susini P., Mcadams S. and Winsberg S.** « Caractérisation perceptive des bruits de véhicules ». Dans *4e Congrès Français d'Acoustique*, p. 543–546, Marseille, (1997).
- [44] **Bezât M.-C.** « Qualification acoustique du typage sonore. application au typage sport ». Rapport de stage DEA-ATIAM. *Paris VI, ENST, Aix-Marseille II, UJF Grenoble I*, (2003).
- [45] **Belda N. and Flindell I.H.** « Automotive interior sound quality : The UK part of a large international comparative study ». Dans *Proceedings of Internoise 2004*, (2004).
- [46] **Fletcher H. and Munson W. A.** « Loudness, its definition, measurement and calculation ». *Journal of the Acoustical Society of America*, 5 : p.82–108, (1933).
- [47] **Courbes d'isotonie de Fletcher et Munson.**
- [48] **Norme ISO 532 : Méthode de calcul du niveau d'isotonie**, (1975).
- [49] **Zwicker E.** « Procedure for calculating loudness of temporally variable sounds ». *Journal of the Acoustical Society of America*, 62(9) : p.675–682, (1977).
- [50] **Moore B.C.J. and Glasberg B.R.** « A revision of Zwicker's loudness model ». *Acta Acustica united with Acustica*, 82 : p.335–345, (1996).
- [51] **Boulet I.** « Sonie des sons impulsionnels : perception, mesures et modèles ». Thèse de doctorat, *Université de la Méditerranée Aix-Marseille II*, (2005).
- [52] **Office of Noise Abatements and Washington Control.** « Annoyance, loudness and measurement of repetitive type impulsive noise sources ». *US Environmental Protection Agency (EPA)*, (1979).
- [53] **Bismarck G.** « Sharpness as an attribute of the timbre of steady sounds ». *Acustica*, 30 : p.159–172, (1974).
- [54] **Terhardt E. and Stoll G.** « Skalierung des Wohlklangs von 17 Umweltschallen und Untersuchung der beteiligten Hörparameter ». *Acustica*, 48 : p.247–253, (1981).
- [55] **Aures von W.** « Ein Berechnungsverfahren der Rauigkeit ». *Acustica*, 58 : p.268–281, (1985).
- [56] **ANSI (1960).** *USA Standard Acoustical Terminology*. american national standards institute.
- [57] **Risset J.-C.** « Paradoxe de hauteur : Le concept de hauteur sonore n'est pas le même pour tout le monde ». *Proceedings of the Seventh International Congress on Acoustics*, Budapest, p. 613–616, (1971).

- [58] **Grey J.M.** « Timbre discrimination in musical patterns ». *Journal of the Acoustical Society of America*, 64 : p.457–472, (1977).
- [59] **Terhardt von E.** « Über akustische rauigkeit und schwankungsstärke ». *Acustica*, 20 : p.215–224, (1968).
- [60] **Pressnitzer D.** « Perception de rugosité psychoacoustique : d'un attribut élémentaire de l'audition à l'écoute musicale ». Thèse de doctorat, *Université Paris VI*, (1998).
- [61] **Plomp R. and Levelt W.J.M.** « Tonal consonance and critical bandwidth ». *Journal of the Acoustical Society of America*, 38 : p.548–560, (1965).
- [62] **Daniel P. and Weber R.** « Psychoacoustical roughness : implementation of an optimized model ». *Acta Acustica united with Acustica*, 83 : p.113–123, (1997).
- [63] **Fastl H.** « Fluctuation strength and temporal masking patterns of amplitude modulated broadband noise ». *Hearing research*, 8 : p.59–69, (1982).
- [64] **Payri F., Torregosa A.J., Broatch A., Marant V. and Beauge I.** « Méthodologie d'étude de la qualité du bruit de combustion d'un moteur diesel d'automobile à partir de l'analyse de sa pression en cylindre ». *Acoustique et Techniques*, 30 : p.25–29, (2002).
- [65] **International Standard ISO 2204-1979(E).** « Acoustics - guide to international standards on the measurement of airborne acoustical noise and evaluation of its effects on human beings ». *Geneva, Switzerland : International Organization for Standardization*, (1979).
- [66] **Définition de son impulsif.** Université de médecine de Rennes. Source Internet : <http://www.med.univ-rennes1.fr/>.
- [67] **Lowet G., Van de Ponsele P., Pauwels S., Van Wayenberge T. and Sas P.** « Development of a metric to quantify diesel engine irregularities ». Dans *Proceedings ISMA 23*, p. 901–906, (1998).
- [68] **Bodden M. and Heinrichs R.** « Diesel sound quality analysis and evaluation ». Dans *Forum Acusticum*, Budapest, (2005).
- [69] **Cohen L.** « Time-frequency analysis ». *Prentice-Hall, Englewood Cliffs, N.J.*, (1995).
- [70] **Gabor D.** « Theory of communication ». *J. Inst. Elec. Eng. London*, 93(26) : p.429–457, (1946).
- [71] **Guillemain P. and Kronland-Martinet R.** « Characterization of acoustic signals through continuous linear time-frequency representations ». *Proceeding of the IEEE : Special issue on time-frequency and time-scale analysis*, 84(4) : p.561–585, april (1996).
- [72] **Delprat N., Escudié B., Guillemain P., Kronland-Martinet R., Tchamitchian P. and Torresani B.** « Asymptotic wavelet and gabor analysis : Extraction of instantaneous frequency ». *IEEE Trans. Inform. Theory*, 38(2) : p.644–664, march (1992).
- [73] **Guillemain P.** « Analyse et modélisation de signaux sonores par les représentations temps-fréquence linéaires ». Thèse de doctorat, *Université d'Aix-Marseille II*, (1994).
- [74] **Démocrite.** Source Internet : <http://mendeleviev.cyberscol.qc.ca/carrefour/rescol99/atome-1.html>.
- [75] **Gabor D.** « Acoustical quanta and the theory of hearing ». *Nature*, 159 (4044) : p.591–594, (1947).
- [76] **Bürck W., Kotowski P. and Lichte H.** « Die lautstarke von knacken, gerauschen und tonen ». *Elektrische Nachrichten-Technik*, 12 : p.326–333, (1935).
- [77] **Feichtinger H. G. and Strohmer T.** « Gabor analysis and algorithms ». Applied and Numerical Harmonic Analysis. Theory and applications. *Birkhäuser Boston Inc.*, (1998).
- [78] **Balian R.** « Un principe d'incertitude fort en théorie du signal ou en mécanique quantique ». *C. R. Acad. Sci. Paris*, II(292) : p.1357–1361, (1981).
- [79] **Low F.** « A passion for physics essay in honor of Geoffrey Chew », Chapitre : Complete sets of wave packets, p. 17–22. *World Scientific, Singapore*, (1985).
- [80] **Bastiaans M.J.** « A sampling theorem for the complex spectrogram and Gabor's expansion of a signal in Gaussian elementary signals ». *Optical Engineering*, 20(4) : p.594–598, July/Aug (1981).
- [81] **A.J.E.M. Janssen.** « The Zak transform : A signal transform for sampled time-continuous signals ». *Philips J. Res.*, 43(1) : p.23–69, (1988).
- [82] **Qiu S.** « The undersampled discrete Gabor transform ». *IEEE Trans. on Signal Processing*, 46(5) : p.1221–1228, may (1998).
- [83] **Fourgeau E. Morlet J., Arens G. and Giard D.** « Wave propagation and sampling theory - part i : Complex signal and scattering in multilayered media ». *Geophysics*, 47 : p.203–221, (1982).

- [84] **Mallat S. G.** « A theory of multiresolution signal decomposition : the wavelet representation ». *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 11 : p.674–693, (1989).
- [85] **Daubechies I.** « Ten lectures on wavelets ». *CBMS-NSF Lecture Notes*, (61), (1992).
- [86] **Matlab Tutorial** : <http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/>.
- [87] **Mallat S. and Hwang W.H.** « Singularity detection and processing with wavelets ». *IEEE Trans. on Information Theory*, 38(2) : p.617–643, (1992).
- [88] **Policar R.** « The wavelet tutorial ». Source Internet : <http://users.rowan.edu/polkar/WAVELETS/WTtutorial.html>.
- [89] **Auger F., Flandrin P., Gonçalves P. and Lemoine O.** « Time-frequency toolbox - Tutorial », (1995-1996). Source Internet : <http://gdr-isis.org/tftb/tutorial/>.
- [90] **Kodera K., Villedary C. and Gendrin R.** « A new method for the numerical analysis of non-stationary signals ». *Physics of the Earth and Planetary Interiors*, 12 : p.142–150, (1976).
- [91] **Flandrin P.** « Temps-Fréquence ». *Traité des Nouvelles Technologies, série Traitement du Signal. Hermès, Paris*, (1993).
- [92] **Torrésani B.** « An overview of wavelet analysis and time-frequency analysis (a minicourse) ». Dans *Self-Similar Systems, Proceedings of the International Workshop" July 30 - August 7, 1998, Dubna, Russia*, p. 9–34, *JINR, E5-99-38*, Dubna, Russia, (1999).
- [93] **Patterson R. D. and Hodsworth J.** « A functional model of neural activity patterns and auditory images ». *Advances in speech, hearing and auditory images. London : JAI Press*, (1990).
- [94] **Fletcher H.** « Auditory patterns ». *Rev. Mod. Phys.*, 12 : p.47–65, (1940).
- [95] **Zwicker E. and Feldtkeller R.** « Psychoacoustique ». *Ed. Masson*, (1981).
- [96] **Aramaki M.** « Analyse-synthèse de sons impulsifs : approches physique et perceptive ». Thèse de doctorat, *Université de la Méditerranée - Aix-Marseille II*, (2003).
- [97] **Codage mp3**. Source Internet : <http://www.macmusic.org/articles/view.php/lang/fr/id/51>.
- [98] **Minidisc**. Source Internet : <http://www.minidisc.org>.
- [99] **Moore B.C.J and Glasberg B.R.** « Suggested formulae for calculating auditory filter bandwidths and excitation patterns ». *J. Acoust. Soc. Am.*, 74 : p.750–753, (1983).
- [100] **Slaney M.** « Auditory toolbox ». Rapport 1998-010. *Interval Research Corporation*, (1998).
- [101] **R.D. Patterson, Allerhand M.H. and Giguère C.** « Time-domain modeling of peripheral auditory processing : A modular architecture and a software platform ». *J. Acoust. Soc. Am.*, (98) : p.1890–1894, (1998).
- [102] **Plack C.J and Moore B.C.J.** « Temporal window shape as a function of frequency and level ». *J. Acoust. Soc. Am*, 87 : p.2178–2187, (1990).
- [103] **Jeong H. and Ih J.** « Implementation of a new algorithm using the stft with variable frequency resolution for the time-frequency auditory model ». *J. Audio Eng. Soc.*, (47) : p.240–251, (1999).
- [104] **Agerkvist F.T.** « A time-frequency auditory model using wavelet packets ». *J. Audio Eng. Soc.*, (44) : p.37–50, (1996).
- [105] **O'Donovan J. and Furlong D.** « Perceptually motivated time-fréquence analysis ». *J. Acoust. Soc. Am.*, 117(1) : p.251–262, January (2005).
- [106] **Moore B.C.J.** « An introduction to the Psychology of Hearing ». *Academic, New York*, (1997).
- [107] **Ajvalasit M. and Giacomini J.** « Analysis of variations in diesel engine idle vibration ». Dans *Proceedings of the Institution of Mechanical Engineers, Part D : Journal of Automobile Engineering*, vol. 217, p. 921–933, (2003).
- [108] **Försth M. and Li H.** « Wavelet multirésolution analysis of spray images from diesel injector ». Dans *Proceedings of PSFVIP-3*, Maui, Hawaii, U.S.A, (2001).
- [109] **Analyse EWD**. Source Internet : <http://www.mee.tcd.ie/mmt/ewd/index.php>.
- [110] **Chiollaz M. and Favre B.** « Caractérisation fine de bruit moteur par analyse temps-fréquence de Wigner-Ville ». *Traitement du Signal*, 8 : p.301–318, (1992).
- [111] **Ren Y., Randall R. B. and Milton B. E.** « Influence of the resonant frequency on the control of knock in diesel engines ». Dans *Proceedings of the Institution of Mechanical Engineers, Part D : Journal of Automobile Engineering*, vol. 213, p. 127–133, (1999).

- [112] **Stancovic L. and Böhme J.** « Time-frequency analysis of multiple resonances in combustion engine signals ». *Signal Processing*, 79 : p.15–28, (1999).
- [113] **Auger F. and Flandrin P.** « Improving the readability of time-frequency and time-scale representations by the method of reassignment ». *IEEE Trans. Signal Processing*, 43(5) : p.1068–1089, (1995).
- [114] **Chassande-Mottin E.** « Méthodes de réallocation dans le plan temps-fréquence pour l'analyse et le traitement de signaux non stationnaires ». Thèse de doctorat, *Université de Cergy-Pontoise*, (1998).
- [115] **Kodera K., Gendrin R. and De Villedary C.** « Analysis of time-varying signals with small bt values ». *IEEE Trans. ASSP*, 26(1) : p.64–76, February (1978).
- [116] **Auger F. and Flandrin P.** « Generalization of the reassignment method to all bilinear time-frequency and time scale representations ». Dans *Proceedings ICASSP*, vol. IV, p. 317–320, (1994).
- [117] **Gribonval R.** « Approximations non-linéaires pour l'analyse des signaux sonores ». Thèse de doctorat, *Université de Paris IX Dauphine*, (1999).
- [118] **Granai L.** « Nonlinear approximation with redundant multi-components dictionaries ». Thèse de doctorat, *Ecole Polytechnique Fédérale de Lausanne*, (2006).
- [119] **Davis G.** « Adaptive nonlinear approximations ». Thèse de doctorat, *New York University*, (1994).
- [120] **Mallat S. G. and Zhang Z.** « Matching pursuits with time-frequency dictionaries ». *IEEE Trans. on Signal Processing*, 41 : p.3397–3415, (1993).
- [121] **Molla S.** « Signaux audiophoniques : modélisation hybride et schéma de codage ». Thèse de doctorat, *Université de Provence (Aix-Marseille I)*, (2003).
- [122] **Jaillet F. and Torrèsani B.** « Remarques sur l'adaptativité des représentations temps-fréquence ». Dans *Proceedings of the GRETSI'03, 19-ème colloque sur la traitement du signal*, (2003).
- [123] **Tropp J.** « Just relax : Convex programming methods for subset selection and sparse approximation ». ICES Report 04-04. *UT-Austin*, February (2004).
- [124] **Davis G., Mallat S. and Zhang Z.** « Adaptive time-frequency decompositions with matching pursuits ». *Optical engineering*, 33(7) : p.2183–2191, July (1994).
- [125] **Chen S. and Donoho D.L.** « Atomic decomposition by basis pursuit ». Techn. Report No. 479. *Dept. Statistics, Univ. Stanford*, (1995).
- [126] **Jost P., Vandergheynst P. and Frossard P.** « Tree-based pursuit ». Technical Report TR-ITS-2004 13. *Swiss Federal Institute of Technology Lausanne (EPFL). Signal Processing Institute*, July (2004).
- [127] **Gribonval R.** « Fast matching pursuit with a multiscale dictionary of gaussian chirps ». *IEEE Transactions on Signal Processing*, 49(5) : p.994–1001, may (2001).
- [128] **Qian S. and Chen D.** « Signal representation via adaptive normalized gaussian functions ». *IEEE Trans. on Signal Processing*, 36(1) : p.1–11, (1994).
- [129] **Pati Y. C., Rezaiifar R. and Krishnaprasad P. S.** « Orthonormal matching pursuit : recursive function approximation with applications to wavelet decomposition ». Dans *Proc. 27th Annu. Asilomar Conf. Signals, Syst. Comput.*, Nov (1993).
- [130] **Baraniuk R. and Jones D.** « Shear madness : New orthonormal bases and frames using chirp functions ». *IEEE Transactions on Signal Processing*, December (1993).
- [131] **Mann S. and Haykin S.** « The chirplet transform : A generalization of Gabor's logon transform ». *Vision Interface '91*, p. 205–212, June 3-7 (1991).
- [132] **Steve Mann and Simon Haykin.** « The chirplet transform : Physical considerations ». *IEEE Trans. Signal Processing*, 43(11) : p.2745–2761, November (1995).
- [133] **A. Bultan.** « A four-parameter atomic decomposition of chirplets ». *IEEE Trans. on Signal Processing*, 41 : p.731–745, (1999).
- [134] **Kronland-Martinet R. and Guillemin P.** « Ridges associated to continuous linear time-frequency representations of asymptotic and transients signals ». *Proc. IEEE : Conf. Time-Frequency and Time-Scale Analysis*, p. 177–180, June (1996).
- [135] **Logiciel Guimauve.** Source Internet <http://webast.ast.obs-mip.fr/people/fbracher/>.
- [136] **Logiciel LastWave.** Source Internet : <http://www.cmap.polytechnique.fr/bacry/LastWave/index.html>.

- [137] **Smith J.O.** « Viewpoints on the history of digital synthesis ». Dans *Proceedings of the International Computer Music Conference (ICMC-91, Montreal)*, Computer Music Association, p. 1–10, October (1991).
- [138] **Chowning J.** « The synthesis of complex audio spectra by means of frequency modulations ». *Journal of the Audio Engineering Society*, 21(7) : p.526–34, september (1973).
- [139] **Schaeffer P.** « Traité des objets musicaux ». *Nouvelle éd.*, (1977).
- [140] **Smith J.O.** « Music applications of digital waveguides ». *STAN-M-39, CCRMA, Music Department, Stanford University*, (1987).
- [141] **Cadoz C., Luciani A. and Florens J.-L.** « CORDIS-ANIMA : a modeling and simulation system for sound and image synthesis - the general formalism ». *Computer Music Journal*, 17(1) : p.19–29, (1993).
- [142] **Orgue.** Source Internet : <http://www.orgue-oloron.org/quest-ce.html>.
- [143] **Telharmonium.** Source Internet : <http://fr.wikipedia.org/wiki/Telharmonium>.
- [144] **Orgue Hammond.** Source Internet : <http://www.organplus.com>.
- [145] **McAulay R.J. and Quatieri Th. F.** « Speech analysis/synthesis based on a sinusoidal representation ». *IEEE Trans. on Acoust., Speech and Signal Proc.*, ASSP-34 : p.744–754, (1986).
- [146] **Serra X.** « Musical signal processing », Chapitre : *Musical sound modeling with sinusoids plus noise*. C. Roads and S. Pope and A. Piccialli and G. De Poli, (1997).
- [147] **Serra X. and Smith J.O.** « Spectral modeling synthesis : A soundanalysis/synthesis system based on a deterministic plus stochastic decomposition ». *Computer Music Journal*, 14(4) : p.12–24, (1990).
- [148] **Rodet X.** « Sinusoidal+residual models for musical sound signals analysis/synthesis ». *Applied Signal Processing*, 4(3) : p.131–141, (1997).
- [149] **Marchand S.** « Musical sound effects in the sas model ». Dans *Proceedings of the Digital Audio Effects (DAFx99) Workshop*, p. 139–142, December (1999).
- [150] **Verma T. and Meng T.** « Time scale modification using a sines+transients+noise signal model ». p. 49–52, (1998).
- [151] **Serra X.** « A system for sound analysis-transformation-synthesis based on a deterministic plus stochastic decomposition ». Thèse de doctorat, *Stanford University*, (1989).
- [152] **Richard F. and Roussarie V.** « Sound design in car passenger compartement. process and tool for the control of engine sound character ». Dans *Proceedings of Journées du design sonore*, october 13-15 (2004).
- [153] **Zadorin L.** « Granular synthesis : an introduction ». *Queensland University of Technology*, (1997).
- [154] **Feldbauer K. and Höldrich R.** « Analysis, resynthesis and interpolation of the interior noise of a car ». Dans *Proceedings of the COST G-6 Conference on Digital Audio effects (DAFX-00)*, Verona, Italy, december 7-9 (2000).
- [155] **Special Acoustique Vibration : « L'analyse modale ».** *Bulletin technique Renault*, novembre 1990. Hors série.
- [156] **Guyader G.** « Modélisations temporelles simplifiées des phénomènes physiques à l'origine du bruit moteur dédiées au portage temps réel et aux applications de design sonore ». Thèse de doctorat, *INSA Lyon*, mai (2003).
- [157] **Guyader G. and Hamzaoui N.** « La synthèse sonore des bruits moteur : de la pression cylindre à l'oreille ». *Acoustique et Technique*, 31, (2002).
- [158] **Hamzaoui N. and Boisson C.** « Calcul du rayonnement acoustique par une méthode intégrale simplifiée ». Rapport final 5951555-5951556, *Convention INSA/INRS*, juin (1996).
- [159] **Roads C.** « Representations of musical signals », Chapitre : *Asynchronous granular synthesis*. Representations of musical signals, G.De Poli and A. Piccialli and C.Roads. *Cambridge, Massachusetts : The MIT Press*, (1991).
- [160] **Truax B.** « Real time granular synthesis with a digital signal processor ». *Computer Music Journal*, 12(2) : p.14–26.
- [161] **Rocha Iturbide M.** « Granular Synthesis ». Thèse de doctorat, *Université de Paris VIII, Saint-Denis*, (1999).

- [162] **Moles A.** « L'évolution actuelle des musiques expérimentales ». *Bulletin du Groupe d'Acoustique Musicale*, 33 : p.22–26, (1968).
- [163] **Green D.** « Temporal auditory acuity ». *Psychological review*, 78 : p.540–551, (1971).
- [164] **Boyer F. and Kronland-Martinet R.** « Granular resynthesis and transformation of sounds through wavelet transform analysis ». Dans *ICMC Proceedings*, p. 51–54, (1989).
- [165] **Truax B.** « Real-time granulation of sampled sounds with the dmx-1000 ». Dans *ICMC Proceedings*, p. 138–145, (1987).
- [166] **Lippe C.** « A musical application of real-time granular sampling using the ircam signal processing workstation ». Dans *Proceedings of the 1994 International Computer Music Conference : ICMA*, (1993).
- [167] **Rodet X., Potard Y. and Barrière J.B.** « The CHANT project : From the synthesis of the singing voice to synthesis in general ». *Computer Music Journal*, 8(3) : p.15–31, (1984).
- [168] **Rodet X. and d'Alessandro C.** « Synthese et analyse-synthese par fonctions d'ondes formantiques ». *Journal Acoustique*, 2 : p.163–169, (1989).
- [169] **Clarke J.M., Manning P.D., Berry R. and Purvis A.** « VOCEL : New implementations of the FOF synthesis method ». Dans *Proceedings of the 14th international Computer Music Conference*, Cologne, (1988).
- [170] **Eckel G. and Rocha Iturbide M.** « The developpment of GiST, a granular synthesis toolkit based on an extention of the FOF generator ». Dans *Proceedings of the 1995 international computer conference ICMA*, Banff Canada, (1995).
- [171] **Truax B.** « Discovering inner complexity : Time-shifting and transposition with a real-time granulation technique ». *Computer Music Journal*, 18 : p.38–48, (1994).
- [172] **MacKey J.** « On the perception of density and stratification in granular sonic textures : An exploratory study ». *Journal of New Music Research*, 13(2) : p.171–186, (1984).
- [173] **Smalley D.** « Spectro-morphology and structuring processes ». *The language of electroacoustic music*. Simon Emmerson, New York : Harwood Academic, (1986).
- [174] **McAdams S. and Bregman A.** « Hearing musical streams ». *Computer music journal*, 3(4) : p.26–44, (1979).
- [175] **Bergeaud F.** « Représentations adaptatives d'images numériques (Matching Pursuit) ». Thèse de doctorat, *Ecole Centrale Paris*, (1995).
- [176] **Cherry E.C.** « Some experiments on the recognition of speech, with one or two ears ». *Journal of the Acoustical Society of America*, 25(5) : p.975–979, (1953).
- [177] **Guillaume P.** « La psychologie de la forme ». *Flammarion*, Paris, (1979).
- [178] **Bregman A. S. and Dannenbring G.** « The effect of continuity on auditory stream segregation ». *Perception Psychophysics*, 13 : p.308–312, (1973).
- [179] **Purwins H., Blankertz B. and Obermayer K.** « Computing auditory perception ». *Computing auditory perception. Organised Sound*, 5 : p.159–171, (2000).
- [180] **Warren R.M and Warren R.P.** « Auditory illusions and confusions ». *Scientific american*, 223 : p.30–36, December (1970).
- [181] **Dowling W.J.** « The perception of interliewed melodies ». *Perception and Psychophysics*, 14 : p.37–40, (1973).
- [182] **U. Wegner.** « Cognitive aspects of amadinda xylophone music from buganda : Inherent patterns reconsidered ». *Ethnomusicology*, 37(2) : p.201–241, (1993).
- [183] **Miller G. A. and Heise G. A.** « The trill threshold ». *Journal of the Acoustical Society of America*, 22(5) : p.637–638, (1950).
- [184] **Schouten J. F.** « On the perception of sound and speech ». Dans *Proceedings of the 4th International Congress on Acoustics*, Copenhagen, vol. 2, p. 201–203, (1962).
- [185] **Grimault N., Michey L., Carlyon R.P. and Collet L.** « Evidence for two pitch encoding mechanisms using a selective auditory training paradigm ». *Perception and Psychophysics*, 64 : p.189–197, (2002).
- [186] **Jones M.R. and Yee W.** « Penser les sons. Psychologie cognitive de l'audition », Chapitre : « L'attention aux événements auditifs : le rôle de l'organisation temporelle ». *Presses Universitaire de France*, (1994).

- [187] **Hirsh I.J.** « Auditory perception of temporal order ». *Journal of the Acoustical Society of America*, 31 : p.759–767, (1959).
- [188] **Miller J.L and Dexter E.R.** « Effects of speaking rate and lexical status on phonetic perception ». *Journal of Experimental Psychology : Human perception and performance*, 14 : p.369–378, (1988).
- [189] **Winckel F.** « Music, sound and sensation : A modern exposition ». (1967).
- [190] **Warren R.M. and Ackroft J.M.** « Two types of auditory sequence perception ». *Perception and Psychophysics*, 20 : p.387–394, (1976).
- [191] **Jones M. R.** « Dynamical attending and responses to time ». *Psychological Review*, 96 : p.459–491, (1989).
- [192] **Courcoux P. and Semenou M.** « Les plans expérimentaux ». *Cours de L'école Nationale d'Ingénieurs des Techniques des Industries Agricoles et Alimentaires*, Unité de Statistique, Septembre (2003).
- [193] **Logiciel Nemrodw.** Source Internet : <http://www.nemrodw.com/>.
- [194] **Shepard R. N.** « The analysis of proximities : multidimensional scaling with an unknown distance function I,II ». *Psychometrika*, 27 : p.125–140, 219–246, (1962).
- [195] **Kruskal J.B.** « Nonmetric multidimensional scaling : a numerical method ». *Psychometrika*, 29 : p.115–129, (1964).
- [196] **Faye P., Brémaud D., Teillet E., Courcoux P., Giboreau A. and Nicod H.** « An alternative to external preference mapping based on consumer perceptive mapping ». *Food Quality and Preference*, 17 : p.604–614, (2006).
- [197] **Mein M.** « Prélocalisation par holographie acoustique des sources de bruit principales sur moteur EW10J4/BE4 adaptation 206 ». (2002). *Note interne PSA*.
- [198] **Maynard D., Williams E. G. and Lee Y.** « I. Theory of generalized holography and the development of NAH ». *The Journal of the Acoustical Society of America*, 78(4) : p.1395–1413, October (1985).